## BUNDESREPUBLIK DEUTSCHLAND



# Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen:

10 2004 007 623.5

Anmeldetag:

17. Februar 2004

Anmelder/Inhaber:

SunGene GmbH & Co KGaA, 06466 Gatersleben/DE

Bezeichnung:

Promotoren zur Expression von Genen in Tagetes

IPC:

C 12 N, A 01 M

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.



München, den 2. September 2004

Deutsches Patent- und Markenamt

Der Präsident
Im Auftrag

Stark

A 9161 03/00 EDV-L Promotoren zur Expression von Genen in Tagetes

#### Beschreibung

Die vorliegende Erfindung betrifft die Verwendung von Promotoren zur Expression, vorzugsweise zur blütenspezifischen Expression von Genen in Pflanzen der Gattung Tagetes, die genetisch veränderten Pflanzen der Gattung Tagetes sowie ein Verfahren zur Herstellung von biosynthetischen Produkten durch Kultivierung der genetisch veränderten Pflanzen.

10

Verschiedene biosynthetische Produkte, wie beispielsweise Feinchemikalien, wie unter anderem Aminosäuren, Vitamine, Carotinoide, aber auch Proteine werden über natürliche Stoffwechselprozesse in Zellen hergestellt und werden in vielen Industriezweigen verwendet, einschließlich der Nahrungsmittel-, Futtermittel-, Kosmetik-, Feed-, Foodund pharmazeutischen Industrie.

5

20

Diese Substanzen, die zusammen als Feinchemikalien/Proteine bezeichnet werden, umfassen unter anderem organische Säuren, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Nukleotide und Nukleoside, Lipide und Fettsäuren, Diole, Kohlenhydrate, aromatische Verbindungen, Vitamine, Carotinoide und Cofaktoren, sowie Proteine und Enzyme. Ihre Produktion im Großmaßstab erfolgt zum Teil mittels biotechnologischer Verfahren unter Verwendung von Mikroorganismen, die entwickelt wurden, um große Mengen der jeweils gewünschten Substanz zu produzieren und sezernieren.

25

Carotinoide werden de novo in Bakterien, Algen, Pilzen und Pflanzen synthetisiert. In den letzten Jahren wird zunehmend versucht, auch Pflanzen als Produktionsorganismen für Feinchemikalien, insbesondere für Vitamine und Carotinoide zu nutzen.

30

Ein natürliches Gemisch aus den Carotinoiden Lutein und Zeaxanthin wird beispielsweise aus den Blüten von Marigold Pflanzen (Tagetes Pflanzen) als sogenanntes Oleoresin extrahiert. Diese Oleoresin findet Anwendung sowohl als Inhaltsstoff von Nahrungsergänzungsmitteln als auch im Feed-Bereich.

35

Lycopin aus Tomaten findet ebenso Anwendung als Nahrungsergänzungsmittel,während Phytoen überwiegend im kosmetischen Bereich verwendet wird.

Ketocarotinoide, also Carotinoide, die mindestens eine Keto-Gruppe enthalten, wie beispielsweise Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin und Adonixanthin sind natürliche Antioxidantien und

40

Pigmente, die von einigen Algen, Pflanzen und Mikroorganismen als Sekundärmetabolite produziert werden.

Aufgrund ihrer farbgebenden Eigenschaften werden die Ketocarotinoide und insbesondere Astaxanthin als Pigmentierhilfsstoffe in der Tierernährung, insbesondere in der Forellen-, Lachs- und Shrimpszucht verwendet.

Ein wirtschaftliches biotechnologisches Verfahren zur Herstellung von natürlichen, biosynthetischen Produkten und insbesondere Carotinoiden ist daher von großer Bedeutung.

WO 0032788 beschreibt einige Carotinoid Biosynthesegene aus Pflanzen der Gattung Tagetes und offenbart, wie genetisch veränderte Pflanzen der Gattung Tagetes hergestellt werden könnten, um in den Petalen verschiedene Carotinoidprofile zu erhalten und damit gezielt bestimmte Carotinoide herzustellen. Dazu ist es nötig, einige Biosynthesegene überzuexprimieren und andere zu unterdrücken.

Zur Überexpression der neu gefundenen Carotinoid-Biosynthesgene in Pflanzen der Gattung Tagetes wird in WO 0032788 der petalenspezifische Promotor der Ketolase aus Adonis vernalis postuliert.

Aufgrund einer Vielzahl möglicher Schwierigkeiten bei der Überexpression bestimmter Gene besteht ein ständiges Bedürfnis, weitere Promotoren zur Verfügung zu stellen, die eine Expression von Genen in Pflanzen der Gattung Tagetes ermöglichen.

Der Erfindung lag daher die Aufgabe zu Grunde, weitere Promotoren zur Verfügung zu stellen, die die Expression von Genen in Pflanzen der Gattung Tagetes ermöglichen.

Demgemäß wurde gefunden, dass sich die Promotoren, ausgewählt aus der Gruppe

A) EPSPS Promotor

10

20

25

30

35

- B) B-Gene Promotor
- C) PDS Promotor und
- D) CHRC Promotor

sehr gut zur Expression von Genen in Pflanzen der Gattung Tagetes eignen, mit der Maßgabe, dass Gene aus Pflanzen der Gattung Tagetes, die in Wildtyppflanzen der Gattung Tagetes von dem jeweiligen Promotor exprimiert werden, ausgenommen sind.

Die Erfindung betrifft daher die Verwendung eines Promotors, ausgewählt aus der Gruppe

- A) EPSPS Promotor
- B) B-Gene Promotor

5

5

20

25

- C) PDS Promotor und
- D) CHRC Promotor

zur Expression von Genen in Pflanzen der Gattung Tagetes, mit der Maßgabe, dass
10 Gene aus Pflanzen der Gattung Tagetes, die in Wildtyppflanzen der Gattung Tagetes
von dem jeweiligen Promotor exprimiert werden, ausgenommen sind.

Benfey et al. (Plant Cell Volume 2, pp. 849-856) beschreiben den EPSPS Promotor aus Petunia als petalenspezifischen Promotor zur Expression von Genen in Petunia hybrida.

Ronen et al. (PNAS Volume 97, Number 20, 11102-11107 beschreiben den B-GENE Promotor aus Tomate als blütenspezifischen Promotor zur Expression von Genen in Tomaten.

Corona et al. (Plant Journal Volume 9 Number 4 pp. 505-512), Mann et al. (Nature Biotechnology Volume 18 pp. 888-892) und Rosati et al. (Plant Journal Volume 24 Number 3 413-419) beschreiben den PDS Promotor aus Tomate als frucht- und blütenspezifischen Promotor zur Expression von Genen in Tomaten und Tabak.

Vishnevetsky et al. (Plant Journal Volume 20 Number 4 pp. 423-431) beschreiben den CHRC Promotor aus Gurke als blütenspezifischen Promotor zur Expression von Genen in Gurke, und weiteren Pflanzen wie z.B. Nelke, Sonnenblume, Tabak.

Es sind weiterhin zahlreiche blütenspezifische Promotoren aus verschiedenen Organismen in der Literatur bekannt. Dabei wurde überraschend festgestellt, dass viele dieser Promotoren in Pflanzen der Gattung Tagetes nicht zur Expression, insbesondere nicht zur blütenspezifischen oder petalenspezifischen Expression von Genen führen.

35 Es war daher überraschend, dass sich die Promotoren, ausgewählt aus der Gruppe

- A) EPSPS Promotor
- B) B-Gene Promotor

C) PDS Promotor und

D) CHRC Promotor

sehr gut zur Expression, insbesondere zur blütenspezifischen und besonderes bevorzugt zur petalenspezifischen Expression von Genen in Pflanzen der Gattung Tagetes eignen.

Unter einem Promotor wird erfindungsgemäß eine Nukleinsäure mit Expressionsaktivität verstanden, also eine Nukleinsäure verstanden, die in funktioneller Verknüpfung mit einer zu exprimierenden Nukleinsäure, im folgenden auch Gen bezeichnet, die Expression, also die Transkription und die Translation dieser Nukleinsäure oder dieses Gens reguliert.

15

10

Unter "Transkription" wird erfindungsgemäß der Prozess verstanden, durch den ausgehend von einer DNA-Matrize ein komplementäres RNA-Molekül hergestellt wird. An diesem Prozess sind Proteine wie die RNA-Polymerase, sogenannte Sigma-Faktoren und transkriptionelle Regulatorproteine beteiligt. Die synthetisierte RNA dient dann als Matrize im Prozess der Translation, der dann zum biosynthetisch aktiven Protein führt.

20 Unter einer "funktionellen Verknüpfung" versteht man in diesem Zusammenhang beispielsweise die sequentielle Anordnung einer der erfindungsgemäßen Promotoren und einer zu exprimierenden Nukleinsäureseguenz und ggf, weiterer regulativer Elemente wie zum Beispiel einem Terminator derart, dass iedes der regulativen Elemente seine Funktion bei der Expression der Nukleinsäuresequenz erfüllen kann. Dazu ist nicht 25 unbedingt eine direkte Verknüpfung im chemischen Sinne erforderlich. Genetische Kontrollsequenzen, wie zum Beispiel Enhancer-Sequenzen, können ihre Funktion auch von weiter entfernten Positionen oder gar von anderen DNA-Molekülen aus auf die Zielsequenz ausüben. Bevorzugt sind Anordnungen, in denen die zu exprimierende Nukleinsäuresequenz oder das zu exprimierende Gen hinter (d.h. am 3'-Ende) der erfindungsgemäßen Promotorsequenz positioniert wird, so dass beide Sequenzen kovalent miteinander verbunden sind. Bevorzugt ist dabei der Abstand zwischen der Promotorsequenz und der zu exprimierende Nukleinsäuresequenz geringer als 200 Basenpaare, besonders bevorzugt kleiner als 100 Basenpaare, ganz besonders bevorzugt kleiner als 50 Basenpaare.

35

Unter "Expressionsaktivität" wird erfindungsgemäß die in einer bestimmten Zeit durch den Promotor gebildete Menge Protein, also die Expressionsrate, verstanden.

15

35

40

5

Unter "spezifischer Expressionsaktivität" wird erfindungsgemäß die in einer bestimmten Zeit durch den Promotor gebildete Menge Protein pro Promotor verstanden.

Bei einer "verursachten Expressionsaktivität" oder "verursachten Expressionsrate" im Bezug auf ein Gen im Vergleich zum Wildtyp wird somit im Vergleich zum Wildtyp die Bildung eines Proteins verursacht, das im Wildtyp so nicht vorhanden war.

Bei einer "erhöhten Expressionsaktivität" oder "erhöhten Expressionsrate" im Bezug auf ein Gen im Vergleich zum Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit die gebildete Menge des Proteins erhöht.

Die Bildungsrate, mit der ein biosynthetsich aktives Protein hergestellt wird, ist ein Produkt aus der Rate der Transkription und der Translation. Beide Raten können erfindungsgemäß beeinflusst werden und damit die Rate der Bildung von Produkten in einem Mikroorganismus beeinflussen.

Die Bezeichnung "dass Gene aus Pflanzen der Gattung Tagetes, die in Wildtyppflanzen der Gattung Tagetes von dem "jeweiligen" Promotor exprimiert werden, ausgenommen sind", bedeutet, dass beispielsweise der EPSPS Promotor aus Pflanzen der Gattuung Tagetes nicht zur Expression von EPSPS-Genen aus Pflanzen der Gattung Tagetes verwendet wird. Dahingegen kann das EPSPS-Gen aus Pflanzen der Gattung Tagetes erfindungsgemäß durch einen B-Gene Promotor, PDS Promotor oder CHRC Promotor aus Pflanzen der Gattung Tagetes eprimiert werden.

25 Unter dem Begriff "Wildtyp" oder "Wildtyppflanze" wird erfindungsgemäß die entsprechende Ausgangspflanze der Gattung Tagetes verstanden.

Je nach Zusammenhang kann unter dem Begriff "Pflanze" die Ausgangspflanze (Wildtyp) oder eine erfindungsgemäße, genetisch veränderte Pflanze der Gattung Tagetes oder beides verstanden werden.

Vorzugsweise wird unter "Wildtyp" für die Erhöhung oder Verursachung der Expressionsaktivität oder Expressionsrate und für die Erhöhung des Gehalts an biosynthetischen Produkten die Pflanze *Tagetes erecta*, insbesondere die Pflanze *Tagetes erecta Hybrid 50011* (WO 02012438) als Referenzorganismus verstanden.

Unter einem "EPSPS Promotor" werden Promotoren verstanden, die natürlicherweise in Organismen, vorzugsweise in Pflanzen, die Genexpression einer Nukleinsäure, kodierend eine 5-Enolpyruvylshikimat-3-phosphatsynthase, regulieren, sowie von diesen Promotorsequenzen durch Substitution, Insertion oder Deletion von Nukleotiden oder

durch Fragmentierung dieser Promotorsequenzen ableitbare Nukleinsäuresequenzen, die noch diese Expressionsaktivität aufweisen und somit funktionelle Äquivalente darstellen.

Diese EPSPS Promotorsequenzen aus anderen Organismen, insbesondere Pflanzen, als den nachstehend angegebenen Promotorsequenzen lassen sich insbesondere durch Homologievergleiche in Datenbanken oder Hybridisierungsstudien mit DNA-Bibliotheken verschiedener Organismen unter Verwendung der nachstehend beschriebenen EPSPS Promotorsequenzen oder den Nukleinsäuren, kodierend eine 5-Enolpyruvylshikimat-3-phosphatsynthase, auffinden.

Vorzugsweise werden dazu die Nukleinsäuren, kodierend eine 5-Enolpyruvylshikimat-3-phosphatsynthase, verwendet, da in der kodierenden Sequenz konservierte Bereiche häufiger sind als in der Promotorsequenz.

Unter einer 5-Enolpyruvylshikimat-3-phosphatsynthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Shikimat-3-Phosphat in 5-Enolpyruvylshikimat-3-Phosphat umzuwandeln.

20 Bevorzugte EPSPS Promotoren enthalten

15

25

35

- A1) die Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 oder
- A2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 1, 2 oder 3 aufweist oder
- A3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 unter stringenten Bedingungen hybridisiert oder
- A4) funktionell äquivalente Fragmente der Sequenzen unter A1), A2) oder A3)

Die Nukleinsäuresequenz SEQ. ID. NO. 1 stellt eine Promotorsequenz der 5-Enolpyruvylshikimat-3-phosphatsynthase (EPSPS) aus Petunia hybrida (AAH19653) dar.

Die Nukleinsäuresequenz SEQ. ID. NO. 2 stellt eine Promotorsequenz der 5-Enolpyruvylshikimat-3-phosphatsynthase (EPSPS) aus Petunia hybrida (M37029) dar.

Die Nukleinsäuresequenz SEQ. ID. NO. 3 stellt eine weitere Promotorsequenz der 5-40 Enolpyruvylshikimat-3-phosphatsynthase (EPSPS) aus Petunia hybrida dar.

10

20

25

7

Die Erfindung betrifft weiterhin EPSPS Promotoren, enthaltend eine von diesen Sequenzen (SEQ. ID. NO. 1, 2 oder 3) durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 1,2 oder 3 aufweist.

Weitere natürliche erfindungsgemäße Beispiele für erfindungsgemäße EPSPS Promotoren lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Identitätsvergleiche der Nukleinsäuresequenzen aus Datenbanken mit den vorstehend beschriebenen

Sequenzen SEQ ID NO: 1, 2 oder 3 leicht auffinden.

Künstliche erfindungsgemäße EPSPS Promotor-Sequenzen lassen sich ausgehend von den Sequenzen SEQ ID NO: 1, 2 oder 3 durch künstliche Variation und Mutation, beispielsweise durch Substitution, Insertion oder Deletion von Nukleotiden leicht auffinden.

Die folgenden Definition und Bedingungen der Identitätsvergleiche und Hybridisierungsbedingungen gelten für alle Nukleinsäuren, also alle Promotoren und Gene der Beschreibung.

Unter dem Begriff "Substitution" ist der Austausch einer oder mehrerer Nukleotide durch ein oder mehrere Nukleotide zu verstehen. "Deletion" ist das Ersetzen eines Nukleotides durch eine direkte Bindung. Insertionen sind Einfügungen von Nukleotiden in die Nukleinsäuresequenz, wobei formal eine direkte Bindung durch ein oder mehrere Nukleotide ersetzt wird.

Unter Identität zwischen zwei Nukleinsäuren wird die Identität der Nukleotide über die jeweils gesamte Nukleinsäurelänge verstanden, insbesondere die Identität die durch Vergleich mit Hilfe der Vector NTI Suite 7.1 Software der Firma Informax (USA) unter Anwendung der Clustal Methode (Higgins DG, Sharp PM. Fast and sensitive multiple sequence alignments on a microcomputer. Comput Appl. Biosci. 1989 Apr;5(2):151-1) unter Einstellung folgender Parameter berechnet wird:

35 Multiple alignment parameter:
Gap opening penalty 10
Gap extension penalty 10
Gap separation penalty range 8
Gap separation penalty off
40 % identity for alignment delay 40

Residue specific gaps off Hydrophilic residue gap off Transition weighing 0

5 Pairwise alignment parameter:

FAST algorithm on

K-tuplesize 1

Gap penalty 3

Window size 5

10 Number of best diagonals 5

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 1 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 1, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 2 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 2, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 3 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 3, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Besonders bevorzugte EPSPS Promotoren weisen mit der jeweiligen Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 eine Identität von mindestens 70%, bevorzugter mindestens 80%, mindestens 90%, mindestens 92%, mindestens 95%, mindestens 96%, mindestens 97%, mindestens 98%, besonders bevorzugt mindestens 99% auf.

Weitere natürliche Beispiele für EPSPS Promotoren lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 1, 2 oder 3, aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

20

25

10

15

20

25

35

9

Ein weiterer Gegenstand der Erfindung betrifft daher EPSPS Promotoren, enthaltend eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. No. 1, 2 oder 3 unter stringenten Bedingungen hybridisiert. Diese Nukleinsäuresequenz umfasst mindestens 10, bevorzugter mehr als 12,15,30,50 oder besonders bevorzugt mehr als 150 Nukleotide.

Unter "hybridisieren" versteht man die Fähigkeit eines Poly- oder Oligonukleotids, unter stringenten Bedingungen an eine nahezu komplementäre Sequenz zu binden, während unter diesen Bedingungen unspezifische Bindungen zwischen nicht-komplementären Partnern unterbleiben. Dazu sollten die Sequenzen vorzugsweise zu 90-100% komplementär sein. Die Eigenschaft komplementärer Sequenzen, spezifisch aneinander binden zu können, macht man sich beispielsweise in der Northern- oder Southern-Blot-Technik oder bei der Primerbindung in PCR oder RT-PCR zunutze.

Die Hybridisierung erfolgt erfinungsgemäß unter stringenten Bedingungen. Solche Hybridisierungsbedingungen sind beispielsweise bei Sambrook, J., Fritsch, E.F., Maniatis, T., in: Molecular Cloning (A Laboratory Manual), 2. Auflage, Cold Spring Harbor Laboratory Press, 1989, Seiten 9.31-9.57 oder in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6 beschrieben:

Unter stringenten Hybridisierungs-Bedingungen werden insbesondere verstanden: Die über Nacht Inkubation bei 42°C in einer Lösung bestehend aus 50 % Formamid, 5 x SSC (750 mM NaCl, 75 mM Tri-Natrium Citrat), 50 mM Natrium Phosphat (ph7,6), 5x Denhardt Lösung, 10% Dextransulfat und 20 g/ml denaturierte, gescheerte Lachsspermien-DNA, gefolgt von einem Waschen der Filter mit 0,1x SSC bei 65°C.

Unter einem "funktionell äquivalenten Fragment" werden für Promotoren Fragmente verstanden die im wesentlichen die gleiche Promotoraktivität aufweisen wie die Ausgangssequenz.

Unter "im wesentlichen gleich" wird eine spezifische Expressionsaktivität verstanden die mindestens 50%, vorzugsweise 60%, bevorzugter 70%, bevorzugter 80%, bevorzugter 90%, besonders bevorzugt 95% der spezifischen Expressionsaktivität der Ausgangssequenz aufweist.

Unter "Fragmente" werden Teilsequenzen der durch Ausführungsform A1), A2) oder A3) beschriebenen EPSPS Promotoren verstanden. Vorzusgweise weisen diese Fragmente mehr als 10, bevorzugter aber mehr als 12,15, 30, 50 oder besonders bevorzugt mehr als 150 zusammenhängende Nukleotide der Nukleinsäuresequenz SEQ.

ID. NO. 1, 2 oder 3 auf.

5

10

15

20

25

35

Besonders bevorzugt ist die Verwendung der Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 als EPSPS Promotor, d.h. zur Expression von Genen in Pflanzen der Gattung Tagetes.

Alle vorstehend erwähnten EPSPS Promotoren sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

Unter einem "B-Gene Promotor" werden Promotoren verstanden, die natürlicherweise in Organismen, vorzugsweise in Pflanzen, die Genexpression einer Nukleinsäure, kodierend eine Lycopin-β-Cyclase, insbesondere eine chromoplastenspezifische Lycopin-β-Cyclase, regulieren, sowie von diesen Promotorsequenzen durch Substitution, Insertion oder Deletion von Nukleotiden oder durch Fragmentierung dieser Promotorsequenzen ableitbare Nukleinsäuresequenzen, die noch diese Expressionsaktivität aufweisen und somit funktionelle Äquivalente darstellen.

Diese B-Gene Promotorsequenzen aus anderen Organismen, insbesondere Pflanzen, als den nachstehend angegebenen Promotorsequenzen lassen sich insbesondere durch Homologievergleiche in Datenbanken oder Hybridisierungsstudien mit DNA-Bibliotheken verschiedener Organismen unter Verwendung der nachstehend beschriebenen B-Gene Promotorsequenzen oder den Nukleinsäuren, kodierend eine Lycopin-β-Cyclase, auffinden.

Vorzugsweise werden dazu die Nukleinsäuren, kodierend eine Lycopin-β-Cyclase, verwendet, da in der kodierenden Sequenz konservierte Bereiche häufiger sind als in der Promotorsequenz.

Unter einer Lycopin- $\beta$ -Cyclase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Lycopin in  $\gamma$ -Carotin und/oder  $\beta$ -Carotin umzuwandeln.

### Bevorzugte B-Gene Promotoren enthalten

- B1) die Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 oder
- B2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 4, 5 oder 6 aufweist oder
- B3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO.4, 5 oder 6 unter stringenten Bedingungen hybridisiert oder
- B4) funktionell äquivalente Fragmente der Sequenzen unter B1), B2) oder B3)

Die Nukleinsäuresequenz SEQ. ID. NO. 4 stellt eine Promotorsequenz der chromoplastenspezifischen Lycopin- $\beta$ -Cyclase (B-Gene) aus Lycopersicon esculentum (AAZ51517) dar.



5

10

Die Nukleinsäuresequenz SEQ. ID. NO. 5 stellt eine Promotorsequenz der chromoplastenspezifischen Lycopin- $\beta$ -Cyclase (B-Gene) aus Lycopersicon esculentum (AAZ51521) dar.

- Die Nukleinsäuresequenz SEQ. ID. NO. 6 stellt eine weitere Promotorsequenz der chromoplastenspezifischen Lycopin-β-Cyclase (B-Gene) aus Lycopersicon esculentum dar.
- Die Erfindung betrifft weiterhin B-Gene Promotoren, enthaltend eine von diesen Sequenzen (SEQ. ID. NO. 4, 5 oder 6) durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 4, 5 oder 6 aufweist.



Weitere natürliche erfindungsgemäße Beispiele für erfindungsgemäße B-Gene Promotoren lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Identitätsvergleiche der Nukleinsäuresequenzen aus Datenbanken mit den vorstehend beschriebenen Sequenzen SEQ ID NO: 4, 5 oder 6 leicht auffinden.

Künstliche erfindungsgemäße B-Gene Promotor-Sequenzen lassen sich ausgehend von den Sequenzen SEQ ID NO: 4, 5 oder 6 durch künstliche Variation und Mutation, beispielsweise durch Substitution, Insertion oder Deletion von Nukleotiden leicht auffinden.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 4 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 4, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 5 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 5, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

**1**5

20

25

5

10

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 6 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 6, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Besonders bevorzugte B-Gene Promotoren weisen mit der jeweiligen Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 eine Identität von mindestens 70%, bevorzugter mindestens 80%, mindestens 90%, mindestens 92%, mindestens 95%, mindestens 96%, mindestens 97%, mindestens 98%, besonders bevorzugt mindestens 99% auf.

Weitere natürliche Beispiele für B-Gene Promotoren lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 4, 5 oder 6 aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.



Ein weiterer Gegenstand der Erfindung betrifft daher B-Gene Promotoren, enthaltend eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. No. 4, 5 oder 6 unter stringenten Bedingungen hybridisiert. Diese Nukleinsäuresequenz umfasst mindestens 10, bevorzugter mehr als 12,15,30,50 oder besonders bevorzugt mehr als 150 Nukleotide.

35

Die Hybridisierungbedingungen sind vorstehend beschrieben.

Unter einem "funktionell äquivalenten Fragment" werden für Promotoren Fragmente verstanden, die im wesentlichen die gleiche Promotoraktivität aufweisen wie die Aus-

gangssequenz.

Unter "im wesentlichen gleich" wird eine spezifische Expressionsaktivität verstanden, die mindestens 50%, vorzugsweise 60%, bevorzugter 70%, bevorzugter 80%, bevorzugter 90%, besonders bevorzugt 95% der spezifischen Expressionsaktivität der Ausgangssequenz aufweist.

Unter "Fragmente" werden Teilsequenzen der durch Ausführungsform B1), B2) oder B3) beschriebenen B-Gene Promotoren verstanden. Vorzusgweise weisen diese Fragmente mehr als 10, bevorzugter aber mehr als 12,15, 30, 50 oder besonders bevorzugts mehr als 150 zusammenhängende Nukleotide der Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 auf.

20

25

5

10

Besonders bevorzugt ist die Verwendung der Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 als B-Gene Promotor, d.h. zur Expression von Genen in Pflanzen der Gattung Tagetes.

Alle vorstehend erwähnten B-Gene Promotoren sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

30

Unter einem "PDS Promotor" werden Promotoren verstanden, die natürlicherweise in Organismen, vorzugsweise in Pflanzen, die Genexpression einer Nukleinsäure, kodierend eine Phytoendesaturase, regulieren, sowie von diesen Promotorsequenzen durch Substitution, Insertion oder Deletion von Nukleotiden oder durch Fragmentierung dieser Promotorsequenzen ableitbare Nukleinsäuresequenzen, die noch diese Expressionsaktivität aufweisen und somit funktionelle Äquivalente darstellen.

35

40

Diese PDS Promotorsequenzen aus anderen Organismen, insbesondere Pflanzen, als den nachstehend angegebenen Promotorsequenzen lassen sich insbesondere durch Homologievergleiche in Datenbanken oder Hybridisierungsstudien mit DNA-Bibliotheken verschiedener Organismen unter Verwendung der nachstehend beschriebenen PDS Promotorsequenzen oder den Nukleinsäuren, kodierend eine Phytoende-

saturase, auffinden.

20

25

35

40

Vorzugsweise werden dazu die Nukleinsäuren, kodierend eine Phytoendesaturase, verwendet, da in der kodierenden Sequenz konservierte Bereiche häufiger sind als in der Promotorsequenz.

Unter einer Phytoendesaturase wird vorzugsweise ein Protein verstanden, das die enzymatische Aktivität aufweist, Phytoen in Phytofluen umzuwandeln.

- 10 Bevorzugte PDS Promotoren enthalten
  - C1) die Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 oder
  - C2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 7, 8, 9 oder 10 aufweist oder
  - C3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 unter stringenten Bedingungen hybridisiert oder
  - C4) funktionell äquivalente Fragmente der Sequenzen unter C1), C2) oder C3)

Die Nukleinsäuresequenz SEQ. ID. NO. 7 stellt eine Promotorsequenz der Phytoendesaturase (PDS) aus Lycopersicon esculentum (U46919) dar.

Die Nukleinsäuresequenz SEQ. ID. NO. 8 stellt eine Promotorsequenz der Phytoendesaturase (PDS) aus Lycopersicon esculentum (X78271) dar.

Die Nukleinsäuresequenz SEQ. ID. NO. 9 stellt eine Promotorsequenz der Phytoendesaturase (PDS) aus Lycopersicon esculentum (X171023) dar.

Die Nukleinsäuresequenz SEQ. ID. NO. 10 stellt eine weitere Promotorsequenz der Phytoendesaturase (PDS) aus Lycopersicon esculentum dar.

Die Erfindung betrifft weiterhin PDS Promotoren, enthaltend eine von diesen Sequenzen (SEQ. ID. NO. 7, 8, 9 oder 10) durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 7, 8, 9 oder 10 aufweist.

Weitere natürliche erfindungsgemäße Beispiele für erfindungsgemäße PDS Promotoren lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Identitätsvergleiche der Nukleinsäuresequenzen aus Da-

10

15

20

25

30

40

15

tenbanken mit den vorstehend beschriebenen Sequenzen SEQ ID NO: 7, 8, 9 oder 10 leicht auffinden.

Künstliche erfindungsgemäße PDS Promotor-Sequenzen lassen sich ausgehend von den Sequenzen SEQ ID NO: 7, 8, 9 oder 10 durch künstliche Variation und Mutation, beispielsweise durch Substitution, Insertion oder Deletion von Nukleotiden leicht auffinden.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 7 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 7, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 8 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 8, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 9 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 9, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 10 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 10, insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Besonders bevorzugte PDS Promotoren weisen mit der jeweiligen Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 eine Identität von mindestens 70%, bevorzugter mindestens 80%, mindestens 90%, mindestens 92%, mindestens 95%, mindestens 96%, mindestens 97%, mindestens 98%, besonders bevorzugt mindestens 99% auf.

Weitere natürliche Beispiele für PDS Promotoren lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 7, 8, 9 oder 10 aus verschiedenen Organismen, deren

20

25

35

40

16

genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

Ein weiterer Gegenstand der Erfindung betrifft daher PDS Promotoren, enthaltend eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. No. 7, 8, 9 oder 10 unter stringenten Bedingungen hybridisiert. Diese Nukleinsäuresequenz umfasst mindestens 10, bevorzugter mehr als 12,15,30,50 oder besonders bevorzugt mehr als 150 Nukleotide.

10 Die Hybridisierungbedingungen sind vorstehend beschrieben.

Unter einem "funktionell äquivalenten Fragment" werden für Promotoren Fragmente verstanden, die im wesentlichen die gleiche Promotoraktivität aufweisen wie die Ausgangssequenz.

Unter "im wesentlichen gleich" wird eine spezifische Expressionsaktivität verstanden, die mindestens 50%, vorzugsweise 60%, bevorzugter 70%, bevorzugter 80%, bevorzugter 90%, besonders bevorzugt 95% der spezifischen Expressionsaktivität der Ausgangssequenz aufweist.

Unter "Fragmente" werden Teilsequenzen der durch Ausführungsform C1), C2) oder C3) beschriebenen PDS Promotoren verstanden. Vorzusgweise weisen diese Fragmente mehr als 10, bevorzugter aber mehr als 12,15, 30, 50 oder besonders bevorzugts mehr als 150 zusammenhängende Nukleotide der Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 auf.

Besonders bevorzugt ist die Verwendung der Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 als PDS Promotor, d.h. zur Expression von Genen in Pflanzen der Gattung Tagetes.

Alle vorstehend erwähnten PDS Promotoren sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

Unter einem "CHRC Promotor" werden Promotoren verstanden, die natürlicherweise in Organismen, vorzugsweise in Pflanzen, die Genexpression einer Nukleinsäure, kodierend ein Chromoplasten-assoziiertes Protein C regulieren, sowie von diesen Promotorsequenzen durch Substitution, Insertion oder Deletion von Nukleotiden oder durch Fragmentierung dieser Promotorsequenzen ableitbare Nukleinsäuresequenzen, die noch diese Expressionsaktivität aufweisen und somit funktionelle Äquivalente darstellen.

Diese CHRC Promotorsequenzen aus anderen Organismen, insbesondere Pflanzen, als den nachstehend angegebenen Promotorsequenzen lassen sich insbesondere durch Homologievergleiche in Datenbanken oder Hybridisierungsstudien mit DNA-Bibliotheken verschiedener Organismen unter Verwendung der nachstehend beschriebenen CHRC Promotorsequenzen oder den Nukleinsäuren kodierend ein Chromoplasten-assoziiertes Protein C auffinden.

Vorzugsweise werden dazu die Nukleinsäuren kodierend ein Chromoplastenassoziiertes Protein C verwendet, da in der kodierenden Sequenz konservierte Bereiche häufiger sind als in der Promotorsequenz.

Bevorzugte CHRC Promotoren enthalten

- D1) die Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13 oder 14 oder
- D2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 11, 12, 13 oder 14 aufweist oder
- D3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13 oder 14 unter stringenten Bedingungen hybridisiert oder
- D4) funktionell äquivalente Fragmente der Sequenzen unter D1), D2) oder D3)

Die Nukleinsäuresequenz SEQ. ID. NO. 11 stellt eine Promotorsequenz des Chromoplasten-assoziiertes Protein C (CHRC) aus Gurke (AAV36416) dar.

Die Nukleinsäuresequenz SEQ. ID. NO. 12 stellt eine weitere Promotorsequenz des Chromoplasten-assoziiertes Protein C (CHRC) aus Gurke dar.

Die Nukleinsäuresequenz SEQ. ID. NO. 13 stellt eine weitere Promotorsequenz des Chromoplasten-assoziiertes Protein C (CHRC) aus Gurke dar.

5

20

25



40

5

25

30

35

40

18

Die Nukleinsäuresequenz SEQ. ID. NO. 14 stellt eine weitere Promotorsequenz des chromoplasten assoziiertes Protein C (CHRC) aus Gurke dar.

Die Erfindung betrifft weiterhin CHRC Promotoren, enthaltend eine von diesen Sequenzen (SEQ. ID. NO. 11, 12, 13, oder 14) durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 11, 12, 13, oder 14 aufweist.

Weitere natürliche erfindungsgemäße Beispiele für erfindungsgemäße CHRC Promotoren lassen sich beispielsweise aus verschiedenen Organismen, deren genomische
Sequenz bekannt ist, durch Identitätsvergleiche der Nukleinsäuresequenzen aus Datenbanken mit den vorstehend beschriebenen Sequenzen SEQ ID NO: 11, 12, 13, oder
14 leicht auffinden.

Künstliche erfindungsgemäße CHRC Promotor-Sequenzen lassen sich ausgehend von den Sequenzen SEQ ID NO: 11, 12, 13, oder 14 durch künstliche Variation und Mutation, beispielsweise durch Substitution, Insertion oder Deletion von Nukleotiden leicht auffinden.

20 Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 11 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 11, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 12 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 12, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 13 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 13, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Unter einer Nukleinsäuresequenz, die eine Identität von mindestens 60 % mit der Sequenz SEQ ID NO: 14 aufweist, wird dementsprechend eine Nukleinsäuresequenz verstanden, die bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 14,

10

20

25

35

19

insbesondere nach obigen Programmlogarithmus, mit obigem Parametersatz eine Identität von mindestens 60 % aufweist.

Besonders bevorzugte CHRC Promotoren weisen mit der jeweiligen Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13, oder 14 eine Identität von mindestens 70%, bevorzugter mindestens 80%, mindestens 90%, mindestens 92%, mindestens 95%, mindestens 96%, mindestens 97%, mindestens 98%, besonders bevorzugt mindestens 99% auf.

Weitere natürliche Beispiele für CHRC Promotoren lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 11, 12, 13, oder 14 aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

Ein weiterer Gegenstand der Erfindung betrifft daher CHCRC Promotoren, enthaltend eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. No. 11, 12, 13, oder 14 unter stringenten Bedingungen hybridisiert. Diese Nukleinsäuresequenz umfasst mindestens 10, bevorzugter mehr als 12,15,30,50 oder besonders bevorzugt mehr als 150 Nukleotide.

Die Hybridisierungbedingungen sind vorstehend beschrieben.

Unter einem "funktionell äquivalenten Fragment" werden für Promotoren Fragmente verstanden, die im wesentlichen die gleiche Promotoraktivität aufweisen wie die Ausgangssequenz.

Unter "im wesentlichen gleich" wird eine spezifische Expressionsaktivität verstanden, die mindestens 50%, vorzugsweise 60%, bevorzugter 70%, bevorzugter 80%, bevorzugter 90%, besonders bevorzugt 95% der spezifischen Expressionsaktivität der Ausgangssequenz aufweist.

Unter "Fragmente" werden Teilsequenzen der durch Ausführungsform D1), D2) oder D3) beschriebenen CHRC Promotoren verstanden. Vorzusgweise weisen diese Fragmente mehr als 10, bevorzugter aber mehr als 12,15, 30, 50 oder besonders bevorzugts mehr als 150 zusammenhängende Nukleotide der Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13, oder 14 auf.

Besonders bevorzugt ist die Verwendung der Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13, oder 14 als CHRC Promotor, d.h. zur Expression von Genen in Pflanzen der

Gattung Tagetes.

Alle vorstehend erwähnten CHRC Promotoren sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

15

20

25

5

10

Mit den erfindungsgemäßen Promotoren lässt sich prinzipiell jedes Gen, also jede Nukleinsäure, kodierend ein Protein, in Pflanzen der Gattung Tagetes exprimieren, insbesondere blütenspezifisch exprimieren, besonders bevorzugt petalenspezifisch exprimieren.

Diese in Pflanzen der Gattung Tagetes zu exprimierenden Gene werden im folgenden auch "Effektgene" genannt.

Bevorzugte Effektgene sind beispielsweise Gene aus dem Biosynthesweg von Geruchsstoffen und Blütenfarben, deren Expression oder erhöhte Expression in Pflanzen der Gattung Tagetes zu einer Veärnderung der Geruchs und/oder der Blütenfarbe von Blüten der Pflanzen der Gattung Tagetes führt.

30

Die Biosynthese von flüchtigen Geruchskomponenten, speziell in Blüten, wurde in den letzten Jahren an verschiedenen Modellorganismen wie Clarkia breweri und Antirhinum majus L. studiert, Flüchtige Geruchskomponenten werden beispielsweise innerhalb des Monoterpen- und Phenylpropan-Stoffwechsels gebildet werden. Im ersten Fall handelt es sich um Linalool; von den Phenylpropanen sind Methyleneugenol, Benzylacetat, Methylbenzoat und Methylsalicat abgeleitet.

Für die Biosynthese von Linalool, (ISo)Methyleigenol, Benzylacetat und Methylsalicinat sind bevorzugte Gene ausgewählz aus der Gruppe Nukleinsäuren kodierend eine Linalool-Synthase (LIS), Nukleinsäuren kodierend eine S-Adenosyl-L-Met:(iso)-Eugenol-O-Methyltransferase (IEMT), Nukleinsäuren kodierend eine Acetyl-CoA-Benzylalkohol-Acetyltransferase und Nukleinsäuren kodierend eine S-Adenosyl-L-Met:Salicylsäure-Methyltransferase (SAMT). Nukleinsäurensequenzen und Proteinsequenzen zu den genannten enzymatischen Aktivitäten sind in Dudareva et al. Plant Cell 8 (1996), 1137-

15

20

25

21

1148; Wang et al. Plant Physiol. 114 (1997), 213-221 und Dudareva et al. Plant J. 14 (1998) 297-304) beschrieben.

Besonders bevorzugte Effektgene sind Gene aus Biosyntheswegen von biosynthetischen Produkten die in Pflanzen der Gattung Tagetes natürlicherweise, d.h. im Wildtyp oder durch genetische Veränderung des Wildtyps hergestellt werden können, insbesondere in Blüten hergestellt werden können, besonders bevorzugt in Petalen hergestellt werden können.

10 Bevorzugte biosynthetische Produkte sind Feinchemikalien.

Der Begriff "Feinchemikalie" ist im Fachgebiet bekannt und beinhaltet Verbidnungen, die von einem Organismus produziert werden und in verschiedenen Industriezweigen Anwendungen finden, wie bspw., jedoch nicht beschränkt auf die pharmazeutische Industrie, die Landwirtschafts-, Kosmetik , Food und Feed-Industrie. Diese Verbindungen umfassen organische Säuren, wie beispielsweise Weinsäure, Itaconsäure und Diaminopimelinsäure, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Purin- und Pyrimidinbasen, Nukleoside und Nukleotide (wie bspw. beschrieben in Kuninaka, A. (1996) Nucleotides and related compounds, S. 561-612, in Biotechnology Bd. 6, Rehm et al., Hrsg. VCH: Weinheim und den darin enthaltenen Zitaten), Lipide, gesättigte und ungesättigte Fettsäuren (bspw. Arachidonsäure), Diole (bspw. Propandiol und Butandiol), Kohlenhydrate (bspw. Hyaluronsäure und Trehalose), aromatische Verbindungen (bspw. aromatische Amine, Vanillin und Indigo), Vitamine, Carotinoide und Cofaktoren (wie beschrieben in Ullmann's Encyclopedia of Industrial Chemistry, Bd. A27, "Vitamins", S. 443-613 (1996) VCH: Weinheim und den darin enthaltenen Zitaten; und Ong, A.S., Niki, E. und Packer, L. (1995) "Nutrition, Lipids, Health and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for Free Radical Research - Asien, abgehalten am 1.-3. Sept. 1994 in Penang, Malysia, AOCS Press (1995)), Enzyme und sämtliche anderen von Gutcho (1983) in Chemicals by Fermentation, Noyes Data Corporation, ISBN: 0818805086 und den darin angegebenen Literaturstellen, beschriebenen Chemikalien). Der Metabolismus und die Verwendungen bestimmter Feinchemikalien sind nachstehend weiter erläutert.

## Aminosäure-Metabolismus und Verwendungen

Die Aminosäuren umfassen die grundlegenden Struktureinheiten sämtlicher Proteine und sind somit für die normalen Zellfunktionen essentiell. Der Begriff "Aminosäure" ist im Fachgebiet bekannt. Die proteinogenen Aminosäuren, von denen es 20 Arten gibt, dienen als Struktureinheiten für Proteine, in denen sie über Peptidbindungen miteinan-

10

20

25

35

40

der verknüpft sind, wohingegen die nicht-proteinogenen Aminosäuren (von denen Hunderte bekannt sind) gewöhnlich nicht in Proteinen vorkommen (siehe Ullmann's Encyclopedia of Industrial Chemistry, Bd. A2, S. 57-97 VCH: Weinheim (1985)). Die Aminosäuren können in der D- oder L-Konfiguration vorliegen, obwohl L-Aminosäuren gewöhnlich der einzige Typ sind, den man in natürlich vorkommenden Proteinen vorfindet. Biosynthese- und Abbauwege von jeder der 20 proteinogenen Aminosäuren sind sowohl bei prokaryotischen als auch eukaryotischen Zellen gut charakterisiert (siehe bspw. Stryer, L. Biochemistry, 3. Auflage, S. 578-590 (1988)). Die "essentiellen" Aminosäuren (Histidin, Isoleucin, Leucin, Lysin, Methionin, Phenylalanin, Threonin, Tryptophan und Valin), so bezeichnet, da sie aufgrund der Komplexität ihrer Biosynthese mit der Ernährung aufgenommen werden müssen, werden durch einfache Biosyntheseswege in die übrigen 11 "nichtessentiellen" Aminosäuren (Alanin, Arginin, Asparagin, Aspartat, Cystein, Glutamat, Glutamin, Glycin, Prolin, Serin und Tyrosin) umgewandelt. Höhere Tiere besitzen die Fähigkeit, einige dieser Aminosäuren zu synthetisieren, jedoch müssen die essentiellen Aminosäuren mit der Nahrung aufgenommen werden, damit eine normale Proteinsynthese stattfindet.

Abgesehen von ihrer Funktion bei der Proteinbiosynthese sind diese Aminosäuren interessante Chemikalien an sich, und man hat entdeckt, daß viele bei verschiedenen Anwendungen in der Nahrungsmittel-, Futter-, Chemie-, Kosmetik-, Landwirtschafts- und pharmazeutischen Industrie zum Einsatz kommen. Lysin ist nicht nur für die Ernährung des Menschen eine wichtige Aminosäure, sondern auch für monogastrische Tiere, wie Geflügel und Schweine. Glutamat wird am häufigsten als Geschmacksadditiv (Mononatriumglutamat, MSG) sowie weithin in der Nahrungsmittelindustrie verwendet, wie auch Aspartat, Phenylalanin, Glycin und Cystein. Glycin, L-Methionin und Tryptophan werden sämtlich in der pharmazeutischen Industrie verwendet. Glutamin, Valin, Leucin, Isoleucin, Histidin, Arginin, Prolin, Serin und Alanin werden in der pharmazeutischen Industrie und der Kosmetikindustrie verwendet. Threonin, Tryptophan und D-/L-Methionin sind weitverbreitete Futtermittelzusätze (Leuchtenberger, W. (1996) Amino acids - technical production and use, S. 466-502 in Rehm et al., (Hrsg.) Biotechnology Bd. 6, Kapitel 14a, VCH: Weinheim). Man hat entdeckt, daß sich diese Aminosäuren außerdem als Vorstufen für die Synthese von synthetischen Aminosäuren und Proteinen, wie N-Acetylcystein, S-Carboxymethyl-L-cystein, (S)-5-Hydroxytryptophan und anderen, in Ullmann's Encyclopedia of Industrial Chemistry, Bd. A2, S. 57-97, VCH, Weinheim, 1985 beschriebenen Substanzen eignen.

Die Biosynthese dieser natürlichen Aminosäuren in Organismen, die sie produzieren können, bspw. Bakterien, ist gut charakterisiert worden (für einen Überblick der bakteriellen Aminosäure-Biosynthese und ihrer Regulation, s. Umbarger, H.E. (1978) Ann. Rev. Biochem. 47: 533 - 606). Glutamat wird durch reduktive Aminierung von  $\alpha$ -

10

5

20

25

30

35

23

Ketoglutarat, einem Zwischenprodukt im Citronensäure-Zyklus, synthetisiert. Glutamin, Prolin und Arginin werden jeweils nacheinander aus Glutamat erzeugt. Die Biosynthese von Serin erfolgt in einem Dreischritt-Verfahren und beginnt mit 3-Phosphoglycerat (einem Zwischenprodukt bei der Glykolyse), und ergibt nach Oxidations-, Transaminierungs- und Hydrolyseschritten diese Aminosäure. Cystein und Glycin werden jeweils aus Serin produziert, und zwar die erstere durch Kondensation von Homocystein mit Serin, und die letztere durch Übertragung des Seitenketten-β-Kohlenstoffatoms auf Tetrahydrofolat, in einer durch Serintranshydroxymethylase katalysierten Reaktion. Phenylalanin und Tyrosin werden aus den Vorstufen des Glycolyse- und Pentosephosphatweges, Erythrose-4-phosphat und Phosphoenolpyruvat in einem 9-Schritt-Biosyntheseweg synthetisiert, der sich nur in den letzten beiden Schritten nach der Synthese von Prephenat unterscheidet. Tryptophan wird ebenfalls aus diesen beiden Ausgangsmolekülen produziert, jedoch erfolgt dessen Synthese in einem 11-Schritt-Weg. Tyrosin läßt sich in einer durch Phenylalaninhydroxylase katalysierten Reaktion auch aus Phenylalanin herstellen. Alanin, Valin und Leucin sind jeweils Biosyntheseprodukte aus Pyruvat, dem Endprodukt der Glykolyse. Aspartat wird aus Oxalacetat, einem Zwischenprodukt des Citratzyklus, gebildet. Asparagin, Methionin, Threonin und Lysin werden jeweils durch Umwandlung von Aspartat produziert. Isoleucin wird aus Threonin gebildet. In einem komplexen 9-Schritt-Weg erfolgt die Bildung von Histidin aus 5-Phosphoribosyl-1-pyrophosphat, einem aktivierten Zucker.

Aminosäuren, deren Menge den Proteinbiosynthesebedarf der Zelle übersteigt, können nicht gespeichert werden, und werden stattdessen abgebaut, so daß Zwischenprodukte für die Haupt-Stoffwechselwege der Zelle bereitgestellt werden (für einen Überblick siehe Stryer, L., Biochemistry, 3. Aufl. Kap. 21 "Amino Acid Degradation and the Urea Cycle"; S 495-516 (1988)). Die Zelle ist zwar in der Lage, ungewünschte Aminosäuren in nützliche Stoffwechsel-Zwischenprodukte umzuwandeln, jedoch ist die Aminosäure-produktion hinsichtlich der Energie, der Vorstufenmoleküle und der für ihre Synthese nötigen Enzyme aufwendig. Es überrascht daher nicht, daß die Aminosäure-Biosynthese durch Feedback-Hemmung reguliert wird, wobei das Vorliegen einer bestimmten Aminosäure ihre eigene Produktion verlangsamt oder ganz beendet (für einen Überblick über den Rückkopplungs-Mechanismus bei Aminosäure-Biosynthesewegen, siehe Stryer, L., Biochemistry, 3. Aufl., Kap. 24, "Biosynthesis of Amino Acids and Heme", S. 575-600 (1988)). Der Ausstoß einer bestimmten Aminosäure wird daher durch die Menge dieser Aminosäure in der Zelle eingeschränkt.

II. Vitamine, Carotinoide, Cofaktoren und Nutrazeutika-Metabolismus sowie Verwendungen

Vitamine, Carotinoide, Cofaktoren und Nutrazeutika umfassen eine weitere Gruppe von Molekülen. Höhere Tiere haben die Fähigkeit verloren, diese zu synthetisieren und müssen sie somit aufnehmen, obwohl sie leicht durch andere Organismen, wie Bakterien, synthetisiert werden. Diese Moleküle sind entweder biologisch aktive Moleküle an sich oder Vorstufen von biologisch aktiven Substanzen, die als Elektronenträger oder Zwischenprodukte bei einer Reihe von Stoffwechselwegen dienen. Diese Verbindungen haben neben ihrem Nährwert auch einen signifikanten industriellen Wert als Farbstoffe, Antioxidantien und Katalysatoren oder andere Verarbeitungs-Hilfsstoffe. (Für einen Überblick über die Struktur, Aktivität und die industriellen Anwendungen dieser Verbindungen siehe bspw. Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", Bd. A27, S. 443-613, VCH: Weinheim, 1996). Der Begriff "Vitamin" ist im Fachgebiet bekannt und umfaßt Nährstoffe, die von einem Organismus für eine normale Funktion benötigt werden, jedoch nicht von diesem Organismus selbst synthetisiert werden können. Die Gruppe der Vitamine kann Cofaktoren und nutrazeutische Verbindungen umfassen. Der Begriff "Cofaktor" umfaßt nicht-proteinartige Verbindungen, die für das Auftreten einer normalen Enzymaktivität nötig sind. Diese Verbindungen können organisch oder anorganisch sein; die erfindungsgemäßen Cofaktor-Moleküle sind vorzugsweise organisch. Der Begriff "Nutrazeutikum" umfaßt Nahrungsmittelzusätze, die bei Pflanzen und Tieren, insbesondere dem Menschen, gesundheitsfördernd sind. Beispiele solcher Moleküle sind Vitamine, Antioxidantien und ebenfalls bestimmte Lipide (z.B. mehrfach ungesättigte Fettsäuren).

Bevorzugte Feinchemikalien oder biosynthetische Produkte, die in Pflanzen der Gattung Tagetes, insbesondere in Petalen der Blüten der Pflanzen der Gattung Tagetes hergestellt werden können, sind Carotinoide, wie beispielsweise Phytoen, Lycopin, Lutein, Zeaxanthin, Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

30 x

35

5

10

15

20

25

Besonders bevorzugte Carotinoide sind Ketocarotinoide, wie beispielsweise Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

Die Biosynthese dieser Moleküle in Organismen, die zu ihrer Produktion befähigt sind, wie Bakterien, ist umfassend charakterisiert worden (Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", Bd. A27, S. 443-613, VCH: Weinheim, 1996, Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley & Sons; Ong, A.S., Niki, E. und Packer, L. (1995) "Nutrition, Lipids, Health and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for free Radical Research - Asien, abgehalten

10

20

25

25

am 1.-3. Sept. 1994 in Penang, Malaysia, AOCS Press, Champaign, IL X, 374 S).

Thiamin (Vitamin B<sub>1</sub>) wird durch chemisches Kuppeln von Pyrimidin und Thiazol-Einheiten gebildet. Riboflavin (Vitamin B<sub>2</sub>) wird aus Guanosin-5'-triphosphat (GTP) und Ribose-5'-phosphat synthetisiert. Riboflavin wiederum wird zur Synthese von Flavinmononukleotid (FMN) und Flavinadenindinukleotid (FAD) eingesetzt. Die Familie von Verbindungen, die gemeinsam als "Vitamin B6" bezeichnet werden (bspw. Pyridoxin, Pyridoxamin, Pyridoxal-5'-phosphat und das kommerziell verwendete Pyridoxinhydrochlorid), sind alle Derivate der gemeinsamen Struktureinheit 5-Hydroxy-6methylpyridin. Panthothenat (Pantothensäure, R-(+)-N-(2,4-Dihydroxy-3,3-dimethyl-1oxobutyl)- $\beta$ -alanin) kann entweder durch chemische Synthese oder durch Fermentation hergestellt werden. Die letzten Schritte bei der Pantothenat-Biosynthese bestehen aus der ATP-getriebenen Kondensation von β-Alanin und Pantoinsäure. Die für die Biosyntheseschritte für die Umwandlung in Pantoinsäure, in  $\beta$ -Alanin und zur Kondensation in Pantothensäure verantwortlichen Enzyme sind bekannt. Die metabolisch aktive Form von Pantothenat ist Coenzym A, dessen Biosynthese über 5 enzymatische Schritte verläuft. Pantothenat, Pyridoxal-5'-phosphat, Cystein und ATP sind die Vorstufen von Coenzym A. Diese Enzyme katalysieren nicht nur die Bildung von Pantothenat, sondern auch die Produktion von (R)-Pantoinsäure, (R)-Pantolacton, (R)-Panthenol (Provitamin B<sub>5</sub>), Pantethein (und seinen Derivaten) und Coenzym A.

Die Biosynthese von Biotin aus dem Vorstufenmolekül Pimeloyl-CoA in Mikroorganismen ist ausführlich untersucht worden, und man hat mehrere der beteiligten Gene identifiziert. Es hat sich herausgestellt, daß viele der entsprechenden Proteine an der Fe-Cluster-Synthese beteiligt sind und zu der Klasse der nifS-Proteine gehören. Die Liponsäure wird von der Octanonsäure abgeleitet und dient als Coenzym beim Energie-Metabolismus, wo sie Bestandteil des Pyruvatdehydrogenasekomplexes und des α-Ketoglutaratdehydrogenasekomplexes wird. Die Folate sind eine Gruppe von Substanzen, die alle von der Folsäure abgeleitet werden, die wiederum von L-Glutaminsäure, p-Aminobenzoesäure und 6-Methylpterin hergeleitet ist. Die Biosynthese der Folsäure und ihrer Derivate, ausgehend von den metabolischen Stoffwechselzwischenprodukten Guanosin-5'-triphosphat (GTP), L-Glutaminsäure und p-Aminobenzoesäure ist in bestimmten Mikroorganismen eingehend untersucht worden.

Corrinoide (wie die Cobalamine und insbesondere Vitamin B<sub>12</sub>) und die Porphyrine gehören zu einer Gruppe von Chemikalien, die sich durch ein Tetrapyrrol-Ringsystem auszeichnen. Die Biosynthese von Vitamin B<sub>12</sub> ist hinreichend komplex, daß sie noch nicht vollständig charakterisiert worden ist, jedoch ist inzwischen ein Großteil der beteiligten Enzyme und Substrate bekannt. Nikotinsäure (Nikotinat) und Nikotinamid sind
 Pyridin-Derivate, die auch als "Niacin" bezeichnet werden. Niacin ist die Vorstufe der

10

20

25

35

40

26

wichtigen Coenzyme NAD (Nikotinamidadenindinukleotid) und NADP (Nikotinamidadenindinukleotidphosphat) und ihrer reduzierten Formen.

Die Produktion dieser Verbindungen im Großmaßstab beruht größtenteils auf zellfreien chemischen Synthesen, obwohl einige dieser Chemikalien ebenfalls durch großangelegte Anzucht von Mikroorganismen produziert worden sind, wie Riboflavin, Vitamin  $B_6$ , Pantothenat und Biotin. Nur Vitamin  $B_{12}$  wird aufgrund der Komplexität seiner Synthese lediglich durch Fermentation produziert. In-vitro-Verfahren erfordern einen erheblichen Aufwand an Materialien und Zeit und häufig an hohen Kosten.

III. Purin-, Pyrimidin-, Nukleosid- und Nukleotid-Metabolismus und Verwendungen

Gene für den Purin- und Pyrimidin-Stoffwechsel und ihre entsprechenden Proteine sind wichtige Ziele für die Therapie von Tumorerkrankungen und Virusinfektionen. Der Begriff "Purin" oder "Pyrimidin" umfaßt stickstoffhaltige Basen, die Bestandteil der Nukleinsäuren, Coenzyme und Nukleotide sind. Der Begriff "Nukleotid" beinhaltet die grundlegenden Struktureinheiten der Nukleinsäuremoleküle, die eine stickstoffhaltige Base, einen Pentose-Zucker (bei RNA ist der Zucker Ribose, bei DNA ist der Zucker D-Desoxyribose) und Phosphorsäure umfassen. Der Begriff "Nukleosid" umfaßt Moleküle, die als Vorstufen von Nukleotiden dienen, die aber im Gegensatz zu den Nukleotiden keine Phosphorsäureeinheit aufweisen. Durch Hemmen der Biosynthese dieser Moleküle oder ihrer Mobilisation zur Bildung von Nukleinsäuremolekülen ist es möglich, die RNA- und DNA-Synthese zu hemmen; wird diese Aktivität zielgerichtet bei kanzerogenen Zellen gehemmt, läßt sich die Teilungs- und Replikations-Fähigkeit von Tumorzellen hemmen.

Es gibt zudem Nukleotide, die keine Nukleinsäuremoleküle bilden, jedoch als Energiespeicher (d.h. AMP) oder als Coenzyme (d.h. FAD und NAD) dienen.

Mehrere Veröffentlichungen haben die Verwendung dieser Chemikalien für diese medizinischen Indikationen beschrieben, wobei der Purin- und/oder Pyrimidin-Metabolismus beeinflußt wird (bspw. Christopherson, R.I. und Lyons, S.D. (1990) "Potent inhibitors of de novo pyrimidine and purine biosynthesis as chemotherapeutic agents", Med. Res. Reviews 10: 505-548). Untersuchungen an Enzymen, die am Purinund Pyrimidin-Metabolismus beteiligt sind, haben sich auf die Entwicklung neuer Medikamente konzentriert, die bspw. als Immunsuppressionsmittel oder Antiproliferantien verwendet werden können (Smith, J.L. "Enzymes in Nucleotide Synthesis" Curr. Opin. Struct. Biol. 5 (1995) 752-757; Biochem. Soc. Transact. 23 (1995) 877-902). Die Purinund Pyrimidinbasen, Nukleoside und Nukleotide haben jedoch auch andere Einsatzmöglichkeiten: als Zwischenprodukte bei der Biosysnthese verschiedener Feinchemikalien (z.B. Thiamin, S-Adenosyl-methionin, Folate oder Riboflavin), als

35

40

27

kalien (z.B. Thiamin, S-Adenosyl-methionin, Folate oder Riboflavin), als Energieträger für die Zelle (bspw. ATP oder GTP) und für Chemikalien selbst, werden gewöhnlich als Geschmacksverstärker verwendet (bspw. IMP oder GMP) oder für viele medizinische Anwendungen (siehe bspw. Kuninaka, A., (1996) "Nucleotides and Related Compounds in Biotechnology Bd. 6, Rehm et al., Hrsg. VCH: Weinheim, S. 561-612). Enzyme, die am Purin-, Pyrimidin-, Nukleosid- oder Nukleotid-Metabolismus beteiligt sind, dienen auch immer stärker als Ziele, gegen die Chemikalien für den Pflanzenschutz, einschließlich Fungiziden, Herbiziden und Insektiziden entwickelt werden.

Der Metabolismus dieser Verbindungen in Bakterien ist charakterisiert worden (für Ü-10 bersichten siehe bspw. Zalkin, H. und Dixon, J.E. (1992) "De novo purin nucleotide biosynthesis" in Progress in Nucleic Acids Research and Molecular biology, Bd. 42, Academic Press, S. 259-287; und Michal, G. (1999) "Nucleotides and Nucleosides"; Kap. 8 in: Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Wiley, New York). Der Purin-Metabolismus, das Objekt intesiver Forschung, ist für das 15 normale Funktionieren der Zelle essentiell. Ein gestörter Purin-Metabolismus in höheren Tieren kann schwere Erkrankungen verursachen, bspw. Gicht. Die Purinnukleotide werden über eine Reihe von Schritten über die Zwischenverbindung Inosin-5'-phosphat (IMP) aus Ribose-5-phosphat synthetisiert, was zur Produktion von Guanosin-5'monophosphat (GMP) oder Adenosin-5'-monophosphat (AMP) führt, aus denen sich 20 die als Nukleotide verwendeten Triphosphatformen leicht herstellen lassen. Diese Verbindungen werden auch als Energiespeicher verwendet, so daß ihr Abbau Energie für viele verschiedene biochemische Prozesse in der Zelle liefert. Die Pyrimidinbiosynthese erfolgt über die Bildung von Uridin-5'-monophosphat (UMP) aus Ribose-5-phosphat. UMP wiederum wird in Cytidin-5'-triphosphat (CTP) umgewandelt. Die Desoxyformen 25 sämtlicher Nukleotide werden in einer Einschritt-Reduktionsreaktion aus der Diphosphat-Riboseform des Nukleotides zur Diphosphat-Desoxyriboseform des Nukleotides hergestellt. Nach der Phosphorylierung können diese Moleküle an der DNA-Synthese teilnehmen.

### IV. Trehalose-Metabolismus und Verwendungen

Trehalose besteht aus zwei Glucosemolekülen, die über α,α-1,1-Bindung miteinander verknüpft sind. Sie wird gewöhnlich in der Nahrungsmittelindustrie als Süßstoff, als Additiv für getrocknete oder gefrorene Nahrungsmittel sowie in Getränken verwendet. Sie wird jedoch auch in der pharmazeutischen Industrie, der Kosmetik- und Biotechnologie-Industrie angewendet (s. bspw. Nishimoto et al., (1998) US-Patent Nr. 5 759 610; Singer, M.A. und Lindquist, S. Trends Biotech. 16 (1998) 460-467; Paiva, C.L.A. und Panek, A.D. Biotech Ann. Rev. 2 (1996) 293-314; und Shiosaka, M. J. Japan 172 (1997) 97-102). Trehalose wird durch Enzyme von vielen Mikroorganismen produziert

10

5

20

25

35

40

28

und auf natürliche Weise in das umgebende Medium abgegeben, aus dem sie durch im Fachgebiet bekannte Verfahren gewonnen werden kann.

Besonders bevorzugte biosynthetische Produkte sind ausgewählt aus der Gruppe organische Säuren, Proteine, Nukleotide und Nukleoside, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Lipide und Fettsäuren, Diole, Kohlehydrate, aromatische Verbindungen, Vitamine und Cofaktoren, Enzyme und Proteine.

Bevorzugte organische Säuren sind Weinsäure, Itaconsäure und Diaminopimelinsäure

Bevorzugte Nukleoside und Nukleotide sind beispielsweise beschrieben in Kuninaka, A. (1996) Nucleotides and related compounds, S. 561-612, in Biotechnology Bd. 6, Rehm et al., Hrsg. VCH: Weinheim und den darin enthaltenen Zitaten.

Bevorzugte biosynthetische Produkte sind weiterhin Lipide, gesättigte und ungesättigte Fettsäuren, wie beispielsweise Arachidonsäure, Diole wie beispielsweise Propandiol und Butandiol, Kohlenhydrate, wie beispielsweise Hyaluronsäure und Trehalose, aromatische Verbindungen, wie beispielsweise aromatische Amine, Vanillin und Indigo, Vitamine und Cofaktoren, wie beispielsweise beschrieben in Ullmann's Encyclopedia of Industrial Chemistry, Bd. A27, "Vitamins", S. 443-613 (1996) VCH: Weinheim und den darin enthaltenen Zitaten; und Ong, A.S., Niki, E. und Packer, L. (1995) "Nutrition, Lipids, Health and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for Free Radical Research - Asien, abgehalten am 1.-3. Sept. 1994 in Penang, Malysia, AOCS Press (1995)), Enzyme Polyketide (Cane et al. (1998) Science 282: 63-68), und sämtliche anderen von Gutcho (1983) in Chemicals by Fermentation, Noyes Data Corporation, ISBN: 0818805086 und den darin angegebenen Literaturstellen, beschriebenen Chemikalien.

Besonders bevorzugte Gene, die mit den erfindungsgemäßen Promotoren in Pflanzen der Gattung Tagetes exprimiert werden sind demnach Gene, ausgewählt sind aus der Gruppe Nukleinsäuren kodierend ein Protein aus den Biosyntheseweg von proteinogenen und nicht-proteinogenen Aminosäuren, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Nukleotiden und Nukleosiden, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von organischen Säuren, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Lipiden und Fettsäuren, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Diolen, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Konhlenhydraten, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von aromatischen Verbindung, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Vitaminen, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Vitaminen, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Carotinoiden, insbesondere Ketocarotinoiden,

Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Cofaktoren und Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Enzymen.

Bevorzugte Feinchemikalien oder biosynthetische Produkte, die in Pflanzen der Gattung Tagetes, insbesondere in Petalen der Blüten der Pflanzen der Gattung Tagetes hergestellt werden können, sind Carotinoide, wie beispielsweise Phytoen, Lycopin, Lutein, Zeaxanthin, Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

10 Besonders bevorzugte Carotinoide sind Ketocarotinoide, wie beispielsweise Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

Ganz besonders bevorzugte Gene, die mit den erfindungsgemäßen Promotoren in Pflanzen der Gattung Tagetes exprimiert werden sind demnach Gene die Proteine aus dem Biosyntheseweg von Carotinoiden kodiern.

Insbesondere bevorzugt sind Gene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine β-Hydroxylase, Nukleinsäuren kodierend eine β-Cyclase, Nukleinsäuren kodierend eine ε-Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat-A-Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.

Unter einer Ketolase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten, β-Ionon-Ring von Carotinoiden eine Keto-Gruppe einzuführen.

Insbesondere wird unter einer Ketolase ein Protein verstanden, das die enzymatische Aktivität aufweist,  $\beta$ -Carotin in Canthaxanthin umzuwandeln.

20

25

Beispiele für Nukleinsäuren, kodierend eine Ketolase, und die entsprechenden Ketolasen, sind beispielsweise Sequenzen aus

Haematoccus pluvialis, insbesondere aus Haematoccus pluvialis Flotow em. Wille (Accession NO: X86782; Nukleinsäure: SEQ ID NO: 15, Protein SEQ ID NO: 16),

Haematoccus pluvialis, NIES-144 (Accession NO: D45881; Nukleinsäure: SEQ ID NO: 17, Protein SEQ ID NO: 18),

10 Agrobacterium aurantiacum (Accession NO: D58420; Nukleinsäure: SEQ ID NO: 19, Protein SEQ ID NO: 20),

Alicaligenes spec. (Accession NO: D58422; Nukleinsäure: SEQ ID NO: 21, Protein SEQ ID NO: 22),

Paracoccus marcusii (Accession NO: Y15112; Nukleinsäure: SEQ ID NO: 23, Protein SEQ ID NO: 24).

Synechocystis sp. Strain PC6803 (Accession NO: NP442491; Nukleinsäure: SEQ ID NO: 25, Protein SEQ ID NO: 26).

Bradyrhizobium sp. (Accession NO: AF218415; Nukleinsäure: SEQ ID NO: 27, Protein SEQ ID NO: 28).

Nostoc sp. Strain PCC7120 (Accession NO: AP003592, BAB74888; Nukleinsäure: SEQ ID NO: 29, Protein SEQ ID NO: 30).

Haematococcus pluvialis

15

35

(Accession NO: AF534876, AAN03484; Nukleinsäure: SEQ ID NO: 31, Protein : SEQ ID NO: 32)

Paracoccus sp. MBIC1143

(Accession NO: D58420, P54972; Nukleinsäure: SEQ ID NO: 33, Protein : SEQ ID NO: 34)

Brevundimonas aurantiaca

(Accession NO: AY166610, AAN86030; Nukleinsäure: SEQ ID NO: 35, Protein : SEQ ID NO: 36)

40 Nodularia spumigena NSOR10

(Accession NO: AY210783, AAO64399; Nukleinsäure: SEQ ID NO: 37, Protein : SEQ ID NO: 38)

Nostoc punctiforme ATCC 29133

5 (Accession NO: NZ\_AABC01000195, ZP\_00111258; Nukleinsäure: SEQ ID NO: 39, Protein: SEQ ID NO: 40)

Nostoc punctiforme ATCC 29133

(Accession NO: NZ\_AABC01000196; Nukleinsäure: SEQ ID NO: 41, Protein : SEQ ID NO: 42)

Deinococcus radiodurans R1

10

20

(Accession NO: E75561, AE001872; Nukleinsäure: SEQ ID NO: 43, Protein : SEQ ID NO: 44),

Synechococcus sp. WH 8102,

Nukleinsäure: Acc.-No. NZ\_AABD01000001, Basenpaar 1,354,725-1,355,528 (SEQ ID NO: 75), Protein: Acc.-No. ZP\_00115639 (SEQ ID NO: 76) (als putatives Protein annotiert),

Unter einer  $\beta$ -Cyclase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, einen endständigen, linearen Rest von Lycopin in einen  $\beta$ -lonon-Ring zu überführen.

Insbesondere wird unter einer  $\beta$ -Cyclase ein Protein verstanden, das die enzymatische Aktivität aufweist,  $\gamma$ -Carotin in  $\beta$ -Carotin umzuwandeln.

Beispiele für  $\beta$ -Cyclase-Gene sind Nukleinsäuren, kodierend eine  $\beta$ -Cyclase aus Tomate (Accession X86452).(Nukleinsäure: SEQ ID NO: 45, Protein: SEQ ID NO: 46), sowie  $\beta$ -Cyclasen der folgenden Accession Nummern:

S66350 lycopene beta-cyclase (EC 5.5.1.-) - tomato

CAA60119	lycopene synthase [Capsicum annuum]
----------	-------------------------------------

S66349 lycopene beta-cyclase (EC 5.5.1.-) - common tobacco

35 CAA57386 lycopene cyclase [Nicotiana tabacum]

AAM21152 lycopene beta-cyclase [Citrus sinensis]

AAD38049 lycopene cyclase [Citrus x paradisi]
AAN86060 lycopene cyclase [Citrus unshiu]

AAF44700 lycopene beta-cyclase [Citrus sinensis]

40 AAK07430 lycopene beta-cyclase [Adonis palaestina]

	AAG10429	beta cyclase [Tagetes erecta]
	AAA81880	lycopene cyclase
	AAB53337	Lycopene beta cyclase
	AAL92175	beta-lycopene cyclase [Sandersonia aurantiaca]
5	· CAA67331	lycopene cyclase [Narcissus pseudonarcissus]
	AAM45381	beta cyclase [Tagetes erecta]
	AAO18661	lycopene beta-cyclase [Zea mays]
	AAG21133	chromoplast-specific lycopene beta-cyclase [Lycopersicon esculentum]
	AAF18989	lycopene beta-cyclase [Daucus carota]
10	ZP_001140	hypothetical protein [Prochlorococcus marinus str. MIT9313]
	ZP_001050	hypothetical protein [Prochlorococcus marinus subsp. pastoris str. CCMP1378]
	ZP_001046	hypothetical protein [Prochlorococcus marinus subsp. pastoris str. CCMP1378]
15	ZP_001134	hypothetical protein [Prochlorococcus marinus str. MIT9313]
	ZP_001150	hypothetical protein [Synechococcus sp. WH 8102]
	AAF10377	lycopene cyclase [Deinococcus radiodurans]
	BAA29250	393aa long hypothetical protein [Pyrococcus horikoshii]
	BAC77673	lycopene beta-monocyclase [marine bacterium P99-3]
20	AAL01999	lycopene cyclase [Xanthobacter sp. Py2]
	ZP_000190	hypothetical protein [Chloroflexus aurantiacus]
	ZP_000941	hypothetical protein [Novosphingobium aromaticivorans]
	AAF78200	lycopene cyclase [Bradyrhizobium sp. ORS278]
	BAB79602	crtY [Pantoea agglomerans pv. milletiae]
25	CAA64855	lycopene cyclase [Streptomyces griseus]
	AAA21262	dycopene cyclase [Pantoea agglomerans]
	C37802	crtY protein - Erwinia uredovora
	BAB79602	crtY [Pantoea agglomerans pv. milletiae]
	AAA64980	lycopene cyclase [Pantoea agglomerans]
30	AAC44851	lycopene cyclase
	BAA09593	Lycopene cyclase [Paracoccus sp. MBIC1143]
	ZP_000941	hypothetical protein [Novosphingobium aromaticivorans]
	CAB56061	lycopene beta-cyclase [Paracoccus marcusii]
	BAA20275	lycopene cyclase [Erythrobacter longus]
35	ZP_000570	hypothetical protein [Thermobifida fusca]
	ZP_000190	hypothetical protein [Chloroflexus aurantiacus]
	AAK07430	lycopene beta-cyclase [Adonis palaestina]
	CAA67331	lycopene cyclase [Narcissus pseudonarcissus]
	AAB53337	Lycopene beta cyclase
40	BAC77673	lycopene beta-monocyclase [marine bacterium P99-3]
		•

Eine besonders bevorzugte  $\beta$ -Cyclase ist weiterhin die chromoplastenspezifische  $\beta$ -Cyclase aus Tomate (AAG21133) (Nukleinsäure: SEQ. ID. No. 49; Protein: SEQ. ID. No. 50)

5

Unter einer Hydroxylase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten,  $\beta$ -Ionon-Ring von Carotinoiden eine Hydroxy-Gruppe einzuführen.

10 Insbesondere wird unter einer Hydroxylase ein Protein verstanden, das die enzymatische Aktivität aufweist, β-Carotin in Zeaxanthin oder Canthaxanthin in Astaxanthin umzuwandeln.

Beispiele für ein Hydroxylase-Gene sind:

eine Nukleinsäure, kodierend eine Hydroxylase aus Haematococcus pluvialis, Accession AX038729, WO 0061764); (Nukleinsäure: SEQ ID NO: 49, Protein: SEQ ID NO: 50),

20 sowie Hydroxylasen der folgenden Accession Nummern:

[emb]CAB55626.1, CAA70427.1, CAA70888.1, CAB55625.1, AF499108\_1, AF315289\_1, AF296158\_1, AAC49443.1, NP\_194300.1, NP\_200070.1, AAG10430.1, CAC06712.1, AAM88619.1, CAC95130.1, AAL80006.1, AF162276\_1, AAO53295.1, AAN85601.1, CRTZ\_ERWHE\_CRTZ\_PANAN\_BAR79605.1, CRTZ\_ALCOR

25 AAN85601.1, CRTZ\_ERWHE, CRTZ\_PANAN, BAB79605.1, CRTZ\_ALCSP, CRTZ\_AGRAU, CAB56060.1, ZP\_00094836.1, AAC44852.1, BAC77670.1, NP\_745389.1, NP\_344225.1, NP\_849490.1, ZP\_00087019.1, NP\_503072.1, NP\_852012.1, NP\_115929.1, ZP\_00013255.1

30

Eine besonders bevorzugte Hydroxylase ist weiterhin die Hydroxylase aus Tomate (Accession Y14810, CrtR-b2) (Nukleinsäure: SEQ ID NO: 51; Protein: SEQ ID NO. 52).

Unter einer HMG-CoA-Reduktase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A in Mevalonat umzuwandeln.

Unter einer (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat in Isopentenyldiphosphat und Dimethylallyldiphosphate umzuwandeln.

35

20

25

35

34

Unter einer 1-Deoxy-D-Xylose-5-Phosphat-Synthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Hydroxyethyl-ThPP und Glycerinaldehyd-3-Phosphat in 1-Deoxy-D-Xylose-5-Phosphat umzuwandeln.

- 5 Unter einer 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, 1-Deoxy-D-Xylose-5-Phosphat in 2-C-methyl-D-erythritol 4-Phosphat umzuwandeln
- Unter einer Isopentenyl-Diphosphat-Δ- Isomerase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Isopentenyl-Diphosphat in Dimethylallylphosphat umzuwandeln.
  - Unter einer Geranyl-Diphosphat-Synthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Isopentenyl-Diphosphat und Dimethylallylphosphat in Geranyl-Diphosphat umzuwandeln.
  - Unter einer Farnesyl-Diphosphat-Synthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, sequentiell 2 Molekülelsopentenyl-Diphosphatmit Dimethylallyl-Diphosphat und dem resultierenden Geranyl-Diphosphat in Farnesyl-Diphosphat umzuwandeln
  - Unter einer Geranyl-Geranyl-Diphosphat-Synthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Farnesyl-Diphosphat und Isopentenyl-Diphosphat in Geranyl-Geranyl-Diphosphat umzuwandeln.
  - Unter einer Phytoen-Synthase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Geranyl-Geranyl-Diphosphat in Phytoen umzuwandeln.
  - Unter einer Phytoen-Desaturase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, Phytoen in Phytofluen und/oder Phytofluen in ζ-Carotin (Zetacarotin) umzuwandeln.
  - Unter einer Zeta-Carotin-Desaturase wird ein Protein verstanden, das die enzymatische Aktivität aufweist,  $\zeta$ -Carotin in Neurosporin und/oder Neurosporin in Lycopin umzuwandeln.
  - Unter einem crtlSO-Proteins wird ein Protein verstanden, das die enzymatische Aktivität aufweist, 7,9,7',9'-tetra-cis-Lycopin in all-trans-Lycopin umzuwandeln.

Unter einem FtsZ-Protein wird ein Protein verstanden, das eine Zellteilungs und Plastidenteilungs-fördernde Wirkung hat und Homologien zu Tubulinproteinen aufweist.

Unter einem MinD -Protein wird ein Protein verstanden, das eine multifunktionele Rolle bei der Zellteilung aufweist. Es ist eine Membran-assoziierte ATPase und kann innerhalb der Zelle eine oszillierende Bewegung von Pol zu Pol zeigen.

Beispiele für HMG-CoA-Reduktase-Gene sind:

5

15

10 Eine Nukleinsäure, kodierend eine HMG-CoA-Reduktase aus Arabidopsis thaliana, Accession NM\_106299; (Nukleinsäure: SEQ ID NO: 53, Protein: SEQ ID NO: 54),

sowie weitere HMG-CoA-Reduktase --Gene aus anderen Organismen mit den folgenden Accession Nummern:

P54961, P54870, P54868, P54869, O02734, P22791, P54873, P54871, P23228, P13704, P54872, Q01581, P17425, P54874, P54839, P14891, P34135, O64966, P29057, P48019, P48020, P12683, P43256, Q9XEL8, P34136, O64967, P29058, P48022, Q41437, P12684, Q00583, Q9XHL5, Q41438, Q9YAS4, O76819, O28538, Q9Y7D2, P54960, O51628, P48021, Q03163, P00347, P14773, Q12577, Q59468, P04035, O24594, P09610, Q58116, O26662, Q01237, Q01559, Q12649, O74164, O59469, P51639, Q10283, O08424, P20715, P13703, P13702, Q96UG4, Q8SQZ9, O15888, Q9TUM4, P93514, Q39628, P93081, P93080, Q944T9, Q40148, Q84MM0, Q84LS3, Q9Z9N4, Q9KLM0

Beispiele für (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase-Gene sind:

Eine Nukleinsäure, kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase aus Arabidopsis thaliana (lytB/ISPH), ACCESSION AY168881, (Nukleinsäure: SEQ ID NO: 55, Protein: SEQ ID NO:56),

sowie weitere (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase –Gene aus anderen Organismen mit den folgenden Accession Nummern:

T04781, AF270978\_1, NP\_485028.1, NP\_442089.1, NP\_681832.1, ZP\_00110421.1, ZP\_00071594.1, ZP\_00114706.1, ISPH\_SYNY3, ZP\_00114087.1, ZP\_00104269.1, AF398145\_1, AF398146\_1, AAD55762.1, AF514843\_1, NP\_622970.1, NP\_348471.1, NP\_562001.1, NP\_223698.1, NP\_781941.1, ZP\_00080042.1, NP\_859669.1, NP\_214191.1, ZP\_00086191.1, ISPH\_VIBCH, NP\_230334.1, NP\_742768.1,

40 NP\_302306.1, ISPH\_MYCLE, NP\_602581.1, ZP\_00026966.1, NP\_520563.1,

10

36

NP\_253247.1, NP\_282047.1, ZP\_00038210.1, ZP\_00064913.1, CAA61555.1, ZP\_00125365.1, ISPH\_ACICA, EAA24703.1, ZP\_00013067.1, ZP\_00029164.1, NP\_790656.1, NP\_217899.1, NP\_641592.1, NP\_636532.1, NP\_719076.1, NP\_660497.1, NP\_422155.1, NP\_715446.1, ZP\_00090692.1, NP\_759496.1, ISPH\_BURPS, ZP\_00129657.1, NP\_215626.1, NP\_335584.1, ZP\_00135016.1, NP\_789585.1, NP\_787770.1, NP\_769647.1, ZP\_00043336.1, NP\_242248.1, ZP\_00008555.1, NP\_246603.1, ZP\_00030951.1, NP\_670994.1, NP\_404120.1, NP\_540376.1, NP\_733653.1, NP\_697503.1, NP\_840730.1, NP\_274828.1, NP\_796916.1, ZP\_00123390.1, NP\_824386.1, NP\_737689.1, ZP\_00021222.1, NP\_757521.1, NP\_390395.1, ZP\_00133322.1, CAD76178.1, NP\_600249.1, NP\_454660.1, NP\_712601.1, NP\_385018.1, NP\_751989.1

Beispiele für 1-Deoxy-D-Xylose-5-Phosphat-Synthase -Gene sind:

Eine Nukleinsäure, kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase aus Lycopersicon esculentum, ACCESSION #AF143812 (Nukleinsäure: SEQ ID NO: 57, Protein: SEQ ID NO: 58),

sowie weitere 1-Deoxy-D-Xylose-5-Phosphat-Synthase -Gene aus anderen Organismen mit den folgenden Accession Nummern: 20 AF143812\_1, DXS\_CAPAN, CAD22530.1, AF182286\_1, NP\_193291.1, T52289, AAC49368.1, AAP14353.1, D71420, DXS\_ORYSA, AF443590\_1, BAB02345.1, CAA09804.2, NP\_850620.1, CAD22155.2, AAM65798.1, NP\_566686.1, CAD22531.1, AAC33513.1, CAC08458.1, AAG10432.1, T08140, AAP14354.1, AF428463\_1, ZP\_00010537.1, NP\_769291.1, AAK59424.1, NP\_107784.1, NP\_697464.1, 25 NP\_540415.1, NP\_196699.1, NP\_384986.1, ZP\_00096461.1, ZP\_00013656.1, NP\_353769.1, BAA83576.1, ZP\_00005919.1, ZP\_00006273.1, NP\_420871.1, AAM48660.1, DXS\_RHOCA, ZP\_00045608.1, ZP\_00031686.1, NP\_841218.1, ZP\_00022174.1, ZP\_00086851.1, NP\_742690.1, NP\_520342.1, ZP\_00082120.1, NP\_790545.1, ZP\_00125266.1, CAC17468.1, NP\_252733.1, ZP\_00092466.1, NP\_439591.1, NP\_414954.1, NP\_752465.1, NP\_622918.1, NP\_286162.1, NP\_836085.1, NP\_706308.1, ZP\_00081148.1, NP\_797065.1, NP\_213598.1, NP\_245469.1, ZP\_00075029.1, NP\_455016.1, NP\_230536.1, NP\_459417.1, NP\_274863.1, NP\_283402.1, NP\_759318.1, NP\_406652.1, DXS\_SYNLE, DXS\_SYNP7, NP\_440409.1, ZP\_00067331.1, ZP\_00122853.1, NP\_717142.1, 35 ZP\_00104889.1, NP\_243645.1, NP\_681412.1, DXS\_SYNEL, NP\_637787.1, DXS\_CHLTE, ZP\_00129863.1, NP\_661241.1, DXS\_XANCP, NP\_470738.1, NP\_484643.1, ZP\_00108360.1, NP\_833890.1, NP\_846629.1, NP\_658213.1,

NP\_642879.1, ZP\_00039479.1, ZP\_00060584.1, ZP\_00041364.1, ZP\_00117779.1,

NP\_299528.1

10

20

25

35

40

Beispiele für 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase-Gene sind:

5 Eine Nukleinsäure, kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase aus Arabidopsis thaliana, ACCESSION #AF148852, (Nukleinsäure: SEQ ID NO: 59, Protein: SEQ ID NO: 60),

sowie weitere 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase-Gene aus anderen Organismen mit den folgenden Accession Nummern:

AF148852, AY084775, AY054682, AY050802, AY045634, AY081453, AY091405, AY098952, AJ242588, AB009053, AY202991, NP\_201085.1, T52570, AF331705\_1, BAB16915.1, AF367205\_1, AF250235\_1, CAC03581.1, CAD22156.1, AF182287\_1, DXR\_MENPI, ZP\_00071219.1, NP\_488391.1, ZP\_00111307.1, DXR\_SYNLE, AAP56260.1, NP\_681831.1, NP\_442113.1, ZP\_00115071.1, ZP\_00105106.1, ZP\_00113484.1, NP\_833540.1, NP\_657789.1, NP\_661031.1, DXR\_BACHD, NP\_833080.1, NP\_845693.1, NP\_562610.1, NP\_623020.1, NP\_810915.1, NP\_243287.1, ZP\_00118743.1, NP\_464842.1, NP\_470690.1, ZP\_00082201.1, NP\_781898.1, ZP\_00123667.1, NP\_348420.1, NP\_604221.1, ZP\_00053349.1, ZP\_00064941.1, NP\_246927.1, NP\_389537.1, ZP\_00102576.1, NP\_519531.1, AF124757\_19, DXR\_ZYMMO, NP\_713472.1, NP\_459225.1, NP\_454827.1, ZP\_00045738.1, NP\_743754.1, DXR\_PSEPK, ZP\_00130352.1, NP\_702530.1, NP\_841744.1, NP\_438967.1, AF514841\_1, NP\_706118.1, ZP\_00125845.1, NP\_404661.1, NP\_285867.1, NP\_240064.1, NP\_414715.1, ZP\_00094058.1, NP\_791365.1, ZP\_00012448.1, ZP\_00015132.1, ZP\_00091545.1, NP\_629822.1, NP\_771495.1, NP\_798691.1, NP\_231885.1, NP\_252340.1, ZP\_00022353.1, NP\_355549.1, NP\_420724.1, ZP\_00085169.1, EAA17616.1, NP\_273242.1, NP\_219574.1, NP\_387094.1, NP\_296721.1, ZP\_00004209.1, NP\_823739.1, NP\_282934.1, BAA77848.1, NP\_660577.1, NP\_760741.1, NP\_641750.1, NP\_636741.1, NP\_829309.1, NP\_298338.1, NP\_444964.1, NP\_717246.1, NP\_224545.1, ZP\_00038451.1, DXR\_KITGR, NP\_778563.1.

Beispiele für Isopentenyl-Diphosphat-Δ-Isomerase-Gene sind:

Eine Nukleinsäure, kodierend eine Isopentenyl-Diphosphat-Δ-Isomerase aus Adonis palaestina clone ApIPI28, (ipiAa1), ACCESSION #AF188060, veröffentlicht durch Cunningham,F.X. Jr. and Gantt,E.: Identification of multi-gene families encoding isopentenyl diphosphate isomerase in plants by heterologous complementation in Escherichia coli, Plant Cell Physiol. 41 (1), 119-123 (2000) (Nukleinsäure: SEQ ID NO: 61, Protein:

SEQ ID NO: 62),

sowie weitere Isopentenyl-Diphosphat-∆-Isomerase–Gene aus anderen Organismen mit den folgenden Accession Nummern:

5

10

Q38929, O48964, Q39472, Q13907, O35586, P58044, O42641, O35760, Q10132, P15496, Q9YB30, Q8YNH4, Q42553, O27997, P50740, O51627, O48965, Q8KFR5, Q39471, Q39664, Q9RVE2, Q01335, Q9HHE4, Q9BXS1, Q9KWF6, Q9CIF5, Q88WB6, Q92BX2, Q8Y7A5, Q8TT35 Q9KK75, Q8NN99, Q8XD58, Q8FE75, Q46822, Q9HP40, P72002, P26173, Q9Z5D3, Q8Z3X9, Q8ZM82, Q9X7Q6, O13504, Q9HFW8, Q8NJL9, Q9UUQ1, Q9NH02, Q9M6K9, Q9M6K5, Q9FXR6, O81691, Q9S7C4, Q8S3L8, Q9M592, Q9M6K3, Q9M6K7, Q9FV48, Q9LLB6, Q9AVJ1, Q9AVG8, Q9M6K6, Q9AVJ5, Q9M6K2, Q9AYS5, Q9M6K8, Q9AVG7, Q8S3L7, Q8W250, Q94IE1, Q9AVI8, Q9AYS6, Q9SAY0, Q9M6K4, Q8GVZ0, Q84RZ8, Q8KZ12, Q8KZ66, Q8FND7, Q88QC9, Q8BFZ6, BAC26382, CAD94476.

5

Beispiele für Geranyl-Diphosphat-Synthase -Gene sind:

Eine Nukleinsäure, kodierend eine Geranyl-Diphosphat-Synthase aus Arabidopsis thaliana, ACCESSION #Y17376, Bouvier,F., Suire,C., d'Harlingue,A., Backhaus,R.A. and Camara,B.; Molecular cloning of geranyl diphosphate synthase and compartmentation of monoterpene synthesis in plant cells, Plant J. 24 (2), 241-252 (2000) (Nukleinsäure: SEQ ID NO: 63, Protein: SEQ ID NO: 64),

sowie weitere Geranyl-Diphosphat-Synthase-Gene aus anderen Organismen mit den folgenden Accession Nummern:

30

35

Q9FT89, Q8LKJ2, Q9FSW8, Q8LKJ3, Q9SBR3, Q9SBR4, Q9FET8, Q8LKJ1, Q84LG1, Q9JK86

Beispiele für Farnesyl-Diphosphat-Synthase-Gene sind:

Eine Nukleinsäure, kodierend eine Farnesyl-Diphosphat-Synthase aus Arabidopsis thaliana (FPS1), ACCESSION #U80605, veröffentlicht durch Cunillera,N., Arro,M., Delourme,D., Karst,F., Boronat,A. und Ferrer,A.: Arabidopsis thaliana contains two differentially expressed farnesyl-diphosphate synthase genes, J. Biol. Chem. 271 (13), 7774-7780 (1996), (Nukleinsäure: SEQ ID NO: 65, Protein: SEQ ID NO:66),

sowie weitere Farnesyl-Diphosphat-Synthase-Gene aus anderen Organismen mit den folgenden Accession Nummern:

P53799, P37268, Q02769, Q09152, P49351, O24241, Q43315, P49352, O24242, P49350, P08836, P14324, P49349, P08524, O66952, Q08291, P54383, Q45220, P57537, Q8K9A0, P22939, P45204, O66126, P55539, Q9SWH9, Q9AVI7, Q9FRX2, Q9AYS7, Q94IE8, Q9FXR9, Q9ZWF6, Q9FXR8, Q9AR37, O50009, Q94IE9, Q8RVK7, Q8RVQ7, O04882, Q93RA8, Q93RB0, Q93RB4, Q93RB5, Q93RB3, Q93RB1, Q93RB2, Q920E5.

10
Beispiele für Geranyl-geranyl-Diphosphat-Synthase -Gene sind:

Eine Nukleinsäure, kodierend eine Geranyl-geranyl-Diphosphat-Synthase aus Sinaps alba, ACCESSION #X98795, veröffentlicht durch Bonk,M., Hoffmann,B., Von Lintig,J., Schledz,M., Al-Babili,S., Hobeika,E., Kleinig,H. and Beyer,P.: Chloroplast import of four carotenoid biosynthetic enzymes in vitro reveals differential fates prior to membrane binding and oligomeric assembly, Eur. J. Biochem. 247 (3), 942-950 (1997), (Nukleinsäure: SEQ ID NO: 67, Protein: SEQ ID NO: 68),

sowie weitere Geranyl-geranyl-Diphosphat-Synthase-Gene aus anderen Organismen mit den folgenden Accession Nummern:

P22873, P34802 ,P56966, P80042, Q42698, Q92236, O95749, Q9WTN0, Q50727, P24322, P39464, Q9FXR3, Q9AYN2, Q9FXR2, Q9AVG6, Q9FRW4, Q9SXZ5, Q9AVJ7, Q9AYN1, Q9AVJ4, Q9FXR7, Q8LSC5, Q9AVJ6, Q8LSC4, Q9AVJ3, Q9SSU0, Q9SXZ6, Q9SST9, Q9AVJ0, Q9AVI9, Q9FRW3, Q9FXR5, Q94IF0, Q9FRX1, Q9K567, Q93RA9, Q93QX8, CAD95619, EAA31459

Beispiele für Phytoen-Synthase-Gene sind:

Eine Nukleinsäure, kodierend eine Phytoen-Synthase aus Erwinia uredovora, ACCES-SION # D90087; veröffentlicht durch Misawa,N., Nakagawa,M., Kobayashi,K., Yamano,S., Izawa,Y.,Nakamura,K. und Harashima,K.: Elucidation of the Erwinia uredovora carotenoid biosynthetic pathway by functional analysis of gene products expressed in Escherichia coli; J. Bacteriol. 172 (12), 6704-6712 (1990), (Nukleinsäure: SEQ ID NO: 69, Protein: SEQ ID NO: 70),

sowie weitere Phytoen-Synthase –Gene aus anderen Organismen mit den folgenden Accession Nummern:

5

15

25

30

35

15

20

25

35

40

CAB39693, BAC69364, AAF10440, CAA45350, BAA20384, AAM72615, BAC09112, CAA48922, P\_001091, CAB84588, AAF41518, CAA48155, AAD38051, AAF33237, AAG10427, AAA34187, BAB73532, CAC19567, AAM62787, CAA55391, AAB65697, AAM45379, CAC27383, AAA32836, AAK07735, BAA84763, P\_000205, AAB60314, P\_001163, P\_000718, AAB71428, AAA34153, AAK07734, CAA42969, CAD76176, CAA68575, P\_000130, P\_001142, CAA47625, CAA85775, BAC14416, CAA79957, BAC76563, P\_000242, P\_000551, AAL02001, AAK15621, CAB94795, AAA91951, P\_000448

10 Beispiele für Phytoen-Desaturase-Gene sind:

Eine Nukleinsäure, kodierend eine Phytoen-Desaturase aus Erwinia uredovora, AC-CESSION # D90087; veröffentlicht durch Misawa,N., Nakagawa,M., Kobayashi,K., Yamano,S., Izawa,Y.,Nakamura,K. und Harashima,K.: Elucidation of the Erwinia uredovora carotenoid biosynthetic pathway by functional analysis of gene products expressed in Escherichia coli; J. Bacteriol. 172 (12), 6704-6712 (1990), (Nukleinsäure: SEQ ID NO: 71, Protein: SEQ ID NO: 72),

sowie weitere Phytoen-Desaturase –Gene aus anderen Organismen mit den folgenden Accession Nummern:

AAL15300, A39597, CAA42573, AAK51545, BAB08179, CAA48195, BAB82461, AAK92625, CAA55392, AAG10426, AAD02489, AAO24235, AAC12846, AAA99519, AAL38046, CAA60479, CAA75094, ZP\_001041, ZP\_001163, CAA39004, CAA44452, ZP\_001142, ZP\_000718, BAB82462, AAM45380, CAB56040, ZP\_001091, BAC09113, AAP79175, AAL80005, AAM72642, AAM72043, ZP\_000745, ZP\_001141, BAC07889, CAD55814, ZP\_001041, CAD27442, CAE00192, ZP\_001163, ZP\_000197, BAA18400, AAG10425, ZP\_001119, AAF13698, 2121278A, AAB35386, AAD02462, BAB68552, CAC85667, AAK51557, CAA12062, AAG51402, AAM63349, AAF85796, BAB74081, AAA91161, CAB56041, AAC48983, AAG14399, CAB65434, BAB73487, ZP\_001117, ZP\_000448, CAB39695, CAD76175, BAC69363, BAA17934, ZP\_000171, AAF65586, ZP\_000748, BAC07074, ZP\_001133, CAA64853, BAB74484, ZP\_001156, AAF23289, AAG28703, AAP09348, AAM71569, BAB69140, ZP\_000130, AAF41516, AAG18866, CAD95940, NP\_656310, AAG10645, ZP\_000276, ZP\_000192, ZP\_000186, AAM94364, EAA31371, ZP\_000612, BAC75676, AAF65582

Beispiele für Zeta-Carotin-Desaturase-Gene sind:

Eine Nukleinsäure, kodierend eine Zeta-Carotin-Desaturase aus Narcissus pseudonarcissus, ACCESSION #AJ224683, veröffentlicht durch Al-Babili,S., Oelschlegel,J. and

Beyer, P.: A cDNA encoding for beta carotene desaturase (Accession No.AJ224683) from Narcissus pseudonarcissus L.. (PGR98-103), Plant Physiol. 117, 719-719 (1998), (Nukleinsäure: SEQ ID NO: 73, Protein: SEQ ID NO: 74),

5 sowie weitere Zeta-Carotin-Desaturase-Gene aus anderen Organismen mit den folgenden Accession Nummern:

Q9R6X4, Q38893, Q9SMJ3, Q9SE20, Q9ZTP4, O49901, P74306, Q9FV46, Q9RCT2, ZDS\_NARPS, BAB68552.1, CAC85667.1, AF372617\_1, ZDS\_TARER, CAD55814.1, CAD27442.1, 2121278A, ZDS\_CAPAN, ZDS\_LYCES, NP\_187138.1, AAM63349.1, ZDS\_ARATH, AAA91161.1, ZDS\_MAIZE, AAG14399.1, NP\_441720.1, NP\_486422.1, ZP\_00111920.1, CAB56041.1, ZP\_00074512.1, ZP\_00116357.1, NP\_681127.1, ZP\_00114185.1, ZP\_00104126.1, CAB65434.1, NP\_662300.1

Beispiele für crtlSO-Gene sind:

Eine Nukleinsäure, kodierend eine crtISO aus Lycopersicon esculentum; ACCESSION #AF416727, veröffentlicht durch Isaacson,T., Ronen,G., Zamir,D. and Hirschberg,J.: Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of beta-carotene and xanthophylls in plants; Plant Cell 14 (2), 333-342 (2002), (Nukleinsäure: SEQ ID NO: 75, Protein: SEQ ID NO: 76),

sowie weitere crtISO –Gene aus anderen Organismen mit den folgenden Accession Nummern:

AAM53952

10

15

20

25

35

Beispiele für FtsZ-Gene sind:

Eine Nukleinsäure, kodierend eine FtsZ aus Tagetes erecta, ACCESSION #AF251346, veröffentlicht durch Moehs, C.P., Tian, L., Osteryoung, K.W. and Dellapenna, D.: Analysis of carotenoid biosynthetic gene expression during marigold petal development Plant Mol. Biol. 45 (3), 281-293 (2001), (Nukleinsäure: SEQ ID NO: 77, Protein: SEQ ID NO: 78),

sowie weitere FtsZ –Gene aus anderen Organismen mit den folgenden Accession Nummern:

CAB89286.1, AF205858\_1, NP\_200339.1, CAB89287.1, CAB41987.1, AAA82068.1, 40 T06774,AF383876\_1, BAC57986.1, CAD22047.1, BAB91150.1, ZP\_00072546.1,

NP\_440816.1, T51092, NP\_683172.1, BAA85116.1, NP\_487898.1, JC4289, BAA82871.1, NP\_781763.1, BAC57987.1, ZP\_00111461.1, T51088, NP\_190843.1, ZP\_00060035.1, NP\_846285.1, AAL07180.1, NP\_243424.1, NP\_833626.1, AAN04561.1, AAN04557.1, CAD22048.1, T51089, NP\_692394.1, NP\_623237.1, NP\_565839.1, T51090, CAA07676.1, NP\_113397.1, T51087, CAC44257.1, E84778, 5 ZP\_00105267.1, BAA82091.1, ZP\_00112790.1, BAA96782.1, NP\_348319.1, NP\_471472.1, ZP\_00115870.1, NP\_465556.1, NP\_389412.1, BAA82090.1, NP\_562681.1, AAM22891.1, NP\_371710.1, NP\_764416.1, CAB95028.1, FTSZ\_STRGR, AF120117\_1, NP\_827300.1, JE0282, NP\_626341.1, AAC45639.1, NP\_785689.1, NP\_336679.1, NP\_738660.1, ZP\_00057764.1, AAC32265.1, 10 NP\_814733.1, FTSZ\_MYCKA, NP\_216666.1, CAA75616.1, NP\_301700.1, NP\_601357.1, ZP\_00046269.1, CAA70158.1, ZP\_00037834.1, NP\_268026.1, FTSZ\_ENTHR, NP\_787643.1, NP\_346105.1, AAC32264.1, JC5548, AAC95440.1, NP\_710793.1, NP\_687509.1, NP\_269594.1, AAC32266.1, NP\_720988.1, NP\_657875.1, ZP\_00094865.1, ZP\_00080499.1, ZP\_00043589.1, JC7087, 15 NP\_660559.1, AAC46069.1, AF179611\_14, AAC44223.1, NP\_404201.1.

Beispiele für MinD -Gene sind:

25

Eine Nukleinsäure, kodierend eine MinD aus Tagetes erecta, ACCESSION #AF251019, veröffentlicht durch Moehs, C.P., Tian, L., Osteryoung, K.W. und Dellapenna, D.: Analysis of carotenoid biosynthetic gene expression during marigold petal development; Plant Mol. Biol. 45 (3), 281-293 (2001), (Nukleinsäure: SEQ ID NO: 79, Protein: SEQ ID NO: 80),

sowie weitere MinD -Gene mit den folgenden Accession Nummern:

NP\_197790.1, BAA90628.1, NP\_038435.1, NP\_045875.1, AAN33031.1, NP\_050910.1, CAB53105.1, NP\_050687.1, NP\_682807.1, NP\_487496.1, ZP\_00111708.1, ZP\_00071109.1, NP\_442592.1, NP\_603083.1, NP\_782631.1, 30 ZP\_00097367.1, ZP\_00104319.1, NP\_294476.1, NP\_622555.1, NP\_563054.1, NP\_347881.1, ZP\_00113908.1, NP\_834154.1, NP\_658480.1, ZP\_00059858.1, NP\_470915.1, NP\_243893.1, NP\_465069.1, ZP\_00116155.1, NP\_390677.1, NP\_692970.1, NP\_298610.1, NP\_207129.1, ZP\_00038874.1, NP\_778791.1, NP\_223033.1, NP\_641561.1, NP\_636499.1, ZP\_00088714.1, NP\_213595.1, 35 NP\_743889.1, NP\_231594.1, ZP\_00085067.1, NP\_797252.1, ZP\_00136593.1, NP\_251934.1, NP\_405629.1, NP\_759144.1, ZP\_00102939.1, NP\_793645.1, NP\_699517.1, NP\_460771.1, NP\_860754.1, NP\_456322.1, NP\_718163.1, NP\_229666.1, NP\_357356.1, NP\_541904.1, NP\_287414.1, NP\_660660.1, ZP\_00128273.1, NP\_103411.1, NP\_785789.1, NP\_715361.1, AF149810\_1, . 40

25

43

NP\_841854.1, NP\_437893.1, ZP\_00022726.1, EAA24844.1, ZP\_00029547.1, NP\_521484.1, NP\_240148.1, NP\_770852.1, AF345908\_2, NP\_777923.1, .
ZP\_00048879.1, NP\_579340.1, NP\_143455.1, NP\_126254.1, NP\_142573.1, NP\_613505.1, NP\_127112.1, NP\_712786.1, NP\_578214.1, NP\_069530.1, NP\_247526.1, AAA85593.1, NP\_212403.1, NP\_782258.1, ZP\_00058694.1, NP\_247137.1, NP\_219149.1, NP\_276946.1, NP\_614522.1, ZP\_00019288.1, CAD78330.1

Die Erfindung betrifft ferner eine genetisch veränderte Pflanze der Gattung Tagetes,
wobei die genetische Veränderung zu einer Erhöhung oder Verursachung der Expressionsrate mindestens eines Gens im Vergleich zum Wildtyp führt und bedingt ist durch die Regulation der Expression dieses Gens in der Pflanze durch die erfindungsgemäßen Promotoren.

Wie vorstehend erwähnt, wird unter "Expressionsaktivität" erfindungsgemäß die in einer bestimmten Zeit durch den Promotor gebildete Menge Protein, also die Expressionsrate verstanden.

Unter "spezifischer Expressionsaktivität" wird erfindungsgemäß die in einer bestimmten Zeit durch den Promotor gebildete Menge Protein pro Promotor verstanden.

Bei einer "verursachten Expressionsaktivität" oder "verursachten Expressionsrate" im Bezug auf ein Gen im Vergleich zum Wildtyp wird somit im Vergleich zum Wildtyp die Bildung eines Proteins verursacht, das im Wildtyp der Pflanze der Gattung Tagetes nicht vorhanden war.

Beispielsweise weisen Wildtyp-Pflanzen der Gattung Tagetes kein Ketolase-Gen auf. Die Regulation der Expression des Ketolase-Gens in der Pflanze durch die erfindungsgemäßen Promotoren fürht somit zu einer Verursachung der Expressionsrate.

Bei einer "erhöhten Expressionsaktivität" oder "erhöhten Expressionsrate" im Bezug auf ein Gen im Vergleich zum Wildtyp wird somit im Vergleich zum Wildtyp der Pflanze der Gattung Tagetes in einer bestimmten Zeit die gebildete Menge des Proteins erhöht.

35 Beispielsweise weisen Wildtyp-Pflanzen der Gattung Tagetes ein Hydroxylase-Gen auf. Die Regulation der Expression des Hydroxylase-Gens in der Pflanze durch die erfindungsgemäßen Promotoren fürht somit zu einer Erhöhung der Expressionsrate.

In einer bevorzugten Ausführungsform der erfindungsgemäßen genetisch veränderten 40 Pflanzen der Gattung Tagetes wird die Regulation der Expression von Genen in der

10

20

35

44

Pflanze durch die erfindungsgemäßen Promotoren dadurch erreicht, dass man

- a) einen oder mehrere erfindungsgemäße Promotoren in das Genom der Pflanze einbringt, so dass die Expression eines oder mehrerer endogenen Gene unter der Kontrolle der eingebrachten erfindungsgemäßen Promotoren erfolgt oder
- b) ein oder mehrere Gene in das Genom der Pflanze einbringt, so dass die Expression eines oder mehrerer der eingebrachten Gene unter der Kontrolle der endogenen, erfindungsgemäßen Promotoren erfolgt oder
- c) ein oder mehrere Nukleinsäurekonstrukte, enthaltend mindestens einen erfindungsgemäßen Promotor und funktionell verknüpft eine oder mehrere zu exprimierende Gene in die Pflanze einbringt.
- In einer bevorzugten Ausführungsform bringt man gemäß Merkmal c) ein oder mehrere Nukleinsäurekonstrukte, enthaltend mindestens einen erfindungsgemäßen Promotor und funktionell verknüpft eine oder mehrere zu exprimierende Gene, in die Pflanze ein. Die Integration der Nukleinsäurekonstrukte in der Pflanze der Gattung Tagetes kann dabei intrachromosomal oder extrachromosomal erfolgen.
- Bevorzugte erfindungsgemäße Promotoren und bevorzugte zu exprimierende Gene (Effektgene) sind vorstehend beschrieben.
- Im folgenden wird exemplarisch die Herstellung der genetisch veränderten Pflanzen der Gattung Tagetes mit erhöhter oder verursachter Expressionsrate eines Effektgens beschrieben.
  - Die Transformation kann bei den Kombinationen von genetischen Veränderungen einzeln oder durch Mehrfachkonstrukte erfolgen.
  - Die Herstellung der transgenen Pflanzen erfolgt vorzugsweise durch Transformation der Ausgangspflanzen, mit einem Nukleinsäurekonstrukt, das mindestens einen der vorstehend beschriebenen, erfindungsgemäßen Promotoren enthält, die mit einem zu exprimierenden Effektgen und gegebenenfalls weiteren Regulationssignalen funktionell verknüpft sind.
  - Diese Nukleinsäurekonstrukte, in denen die erfindungsgemäßen Promotoren und Effektgene funktionell verknüpft sind, werden im folgenden auch Expressionskassetten genannt.

15

20

25

35

45

Die Expressionskassetten können weitere Regulationssignale enthalten, also regulative Nukleinsäuresequenzen, welche die Expression der Effektgene in der Wirtszelle steuern. Gemäß einer bevorzugten Ausführungsform umfasst eine Expressionskassette stromaufwärts, d.h. am 5'-Ende der kodierenden Sequenz, mindestens einen erfindungsgemäßen Promotor und stromabwärts, d.h. am 3'-Ende, ein Polyadenylierungssignal und gegebenenfalls weitere regulatorische Elemente, welche mit der dazwischenliegenden kodierenden Sequenz des Effektgens für mindestens eines der vorstehend beschriebenen Gene operativ verknüpft sind.

10 Unter einer operativen Verknüpfung versteht man die sequenzielle Anordnung von Promotor, kodierender Sequenz, Terminator und ggf. weiterer regulativer Elemente derart, das jedes der regulativen Elemente seine Funktion bei der Expression der kodierenden Sequenz bestimmungsgemäß erfüllen kann.

Im folgenden werden beispielhaft die bevorzugten Nukleinsäurekonstrukte, Expressionskassetten und Vektoren für Pflanzen und Verfahren zur Herstellung von transgenen Pflanzen, sowie die transgenen Pflanzen der Gattung Tagetes selbst beschrieben.

Die zur operativen Verknüpfung bevorzugten, aber nicht darauf beschränkten Sequenzen, sind Targeting-Sequenzen zur Gewährleistung der subzellulären Lokalisation im Apoplasten, in der Vakuole, in Plastiden, im Mitochondrium, im Endoplasmatischen Retikulum (ER), im Zellkern, in Ölkörperchen oder anderen Kompartimenten und Translationsverstärkern wie die 5'-Führungssequenz aus dem Tabak-Mosaik-Virus (Gallie et al., Nucl. Acids Res. 15 (1987), 8693 -8711).

Die Herstellung einer Expressionskassette erfolgt vorzugsweise durch Fusion mindestens eines erfindungsgemäßen Promotors mit mindestens einem Gen, vorzugsweise mit einem der vorstehend beschriebenen Effektgene, und vorzugsweise einer zwischen Promotor und Nukleinsäure-Sequenz inserierten Nukleinsäure, die für ein plastidenspezifisches Transitpeptid kodiert, sowie einem Polyadenylierungssignal nach gängigen Rekombinations- und Klonierungstechniken, wie sie beispielsweise in T. Maniatis, E.F. Fritsch und J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) sowie in T.J. Silhavy, M.L. Berman und L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) und in Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley-Interscience (1987), beschrieben sind.

Die vorzugsweise insertierte Nukleinsäuren, kodierend ein plastidäres Transitpeptid, gewährleisten die Lokalisation in Plastiden und insbesondere in Chromoplasten.

Es können auch Expressionskassetten verwendet werden, deren Nukleinsäure—Sequenz für ein Effektgen-Produkt-Fusionsprotein kodiert, wobei ein Teil des Fusionsproteins ein Transitpeptid ist, das die Translokation des Polypeptides steuert. Bevorzugt sind für die Chromoplasten spezifische Transitpeptide, welche nach Translokation der Effektgene in die Chromoplasten vom Effektgenprodukt—Teil enzymatisch abgespalten werden.

10

5

Insbesondere bevorzugt ist das Transitpeptid, das von der plastidären *Nicotiana taba-cum* Transketolase oder einem anderen Transitpeptid (z.B. dem Transitpeptid der kleinen Untereinheit der Rubisco (rbcS) oder der Ferredoxin NADP Oxidoreduktase als auch der Isopentenylpyrophosphat Isomerase-2 oder dessen funktionellem Äquivalent abgeleitet ist.

15

Besonders bevorzugt sind Nukleinsäure-Sequenzen von drei Kassetten des Plastiden-Transitpeptids der plastidären Transketolase aus Tabak in drei Leserastern als Kpnl/BamHl Fragmente mit einem ATG-Codon in der Ncol Schnittstelle:

20

25

pTP09

30

40

pTP10

Kpnl\_GGTACCATGGCGTCTTCTTCTCTCACTCTCTCAAGCTATCCTCTCTC
GTTCTGTCCCTCGCCATGGCTCTGCCTCTTCTTCTCAACTTTCCCCTTCTTCTCTCACTTTTTCCGGCCTTAAATCCAATCCCAATATCACCACCTCCCGCCGCCG35 TACTCCTTCCTCCGCCGCCGCCGCCGCCGTCGTAAGGTCACCGGCGATTCGTGCCTCAGCTGCAACCGAAACCATAGAGAAAACTGAGACTGCGCTGGATCC\_BamHI

pTP11

10

15

35

47

Weitere Beispiele für ein plastidäres Transitpeptid sind das Transitpeptid der plastidären Isopentenyl-pyrophosphat Isomerase-2 (IPP-2) aus Arabisopsis thaliana und das Transitpeptid der kleinen Untereinheit der Ribulosebisphosphat Carboxylase (rbcS) aus Erbse (Guerineau, F, Woolston, S, Brooks, L, Mullineaux, P (1988) An expression cassette for targeting foreign proteins into the chloroplasts. Nucl. Acids Res. 16: 11380).

Die erfindungsgemäßen Nukleinsäuren können synthetisch hergestellt oder natürlich gewonnen sein oder eine Mischung aus synthetischen und natürlichen Nukleinsäure-Bestandteilen enthalten, sowie aus verschiedenen heterologen Genabschnitten verschiedener Organismen bestehen.

Bevorzugt sind, wie vorstehend beschrieben, synthetische Nukleotid-Sequenzen mit Kodons, die von Pflanzen bevorzugt werden. Diese von Pflanzen bevorzugten Kodons können aus Kodons mit der höchsten Proteinhäufigkeit bestimmt werden, die in den meisten interessanten Pflanzenspezies exprimiert werden.

Bei der Präparation einer Expressionskassette können verschiedene DNA-Fragmente manipuliert werden, um eine Nukleotid-Sequenz zu erhalten, die zweckmäßigerweise in der korrekten Richtung liest und die mit einem korrekten Leseraster ausgestattet ist. Für die Verbindung der DNA-Fragmente miteinander können an die Fragmente Adaptoren oder Linker angesetzt werden.

Zweckmäßigerweise können die Promotor- und die Terminator-Regionen in Transkriptionsrichtung mit einem Linker oder Polylinker, der eine oder mehrere Restriktionsstellen für die Insertion dieser Sequenz enthält, versehen werden. In der Regel hat der Linker 1 bis 10, meistens 1 bis 8, vorzugsweise 2 bis 6 Restriktionsstellen. Im allgemeinen hat der Linker innerhalb der regulatorischen Bereiche eine Größe von weniger als 100 bp, häufig weniger als 60 bp, mindestens jedoch 5 bp. Der Promotor kann sowohl nativ bzw. homolog als auch fremdartig bzw. heterolog zur Wirtspflanze sein. Die Expressionskassette beinhaltet vorzugsweise in der 5'-3'-Transkriptionsrichtung den Promotor, eine kodierende Nukleinsäuresequenz oder ein Nukleinsäurekonstrukt und eine Region für die transkriptionale Termination. Verschiedene Terminationsbereiche

sind gegeneinander beliebig austauschbar.

Beispiele für einen Terminator sind der 35S-Terminator (Guerineau et al. (1988) Nucl Acids Res. 16: 11380), der nos Terminator (Depicker A, Stachel S, Dhaese P, Zambryski P, Goodman HM. Nopaline synthase: transcript mapping and DNA sequence. J Mol Appl Genet. 1982;1(6):561-73) oder der ocs Terminator (Gielen, J, de Beuckeleer, M, Seurinck, J, Debroek, H, de Greve, H, Lemmers, M, van Montagu, M, Schell, J (1984) The complete sequence of the TL-DNA of the Agrobacterium tumefaciens plasmid pTiAch5. EMBO J. 3: 835-846).

10

5

Ferner können Manipulationen, die passende Restriktionsschnittstellen bereitstellen oder die überflüssige DNA oder Restriktionsschnittstellen entfernen, eingesetzt werden. Wo Insertionen, Deletionen oder Substitutionen wie z.B. Transitionen und Transversionen in Frage kommen, können *in vitro*-Mutagenese, "primer-repair", Restriktion oder Ligation verwendet werden.

15

Bei geeigneten Manipulationen, wie z.B. Restriktion, "chewing-back" oder Auffüllen von Überhängen für "bluntends", können komplementäre Enden der Fragmente für die Ligation zur Verfügung gestellt werden.

20

Bevorzugte Polyadenylierungssignale sind pflanzliche Polyadenylierungssignale, vorzugsweise solche, die im wesentlichen T-DNA-Polyadenylierungssignale aus Agrobacterium tumefaciens, insbesondere des Gens 3 der T-DNA (Octopin Synthase) des Ti-Plasmids pTiACH5 entsprechen (Gielen et al., EMBO J. 3 (1984), 835 ff) oder funktionelle Äquivalente.

25

Die Übertragung von Fremdgenen in das Genom einer Pflanze wird als Transformation bezeichnet.

**---**),

Dazu können an sich bekannte Methoden zur Transformation und Regeneration von Pflanzen aus Pflanzengeweben oder Pflanzenzellen zur transienten oder stabilen Transformation genutzt werden.

35

40

Geeignete Methoden zur Transformation von Pflanzen sind die Protoplastentransformation durch Polyethylenglykol-induzierte DNA-Aufnahme, das biolistische Verfahren mit der Genkanone – die sogenannte "particle bombardment" Methode, die Elektroporation, die Inkubation trockener Embryonen in DNA-haltiger Lösung, die Mikroinjektion und der, vorstehend beschriebene, durch *Agrobacterium* vermittelte Gentransfer. Die genannten Verfahren sind beispielsweise in B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von

15

20

25

35

40

49

S.D. Kung und R. Wu, Academic Press (1993), 128-143 sowie in Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225) beschrieben.

Vorzugsweise wird das zu exprimierende Konstrukt in einen Vektor kloniert, der geeignet ist, *Agrobacterium tumefaciens* zu transformieren, beispielsweise pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984), 8711) oder besonders bevorzugt pSUN2, pSUN3, pSUN4 oder pSUN5 (WO 02/00900).

Mit einem Expressionsplasmid transformierte Agrobakterien können in bekannter Weise zur Transformation von Pflanzen verwendet werden, z.B. indem verwundete Blätter
oder Blattstücke in einer Agrobakterienlösung gebadet und anschließend in geeigneten
Medien kultiviert werden.

Zur bevorzugten Herstellung von genetisch veränderten Pflanzen, im folgenden auch transgene Pflanzen bezeichnet, wird die fusionierte Expressionskassette in einen Vektor, beispielsweise pBin19 oder insbesondere pSUN5 und pSUN3 kloniert, der geeignet ist, in *Agrobacterium tumefaciens* transformiert zu werden. Mit einem solchen Vektor transformierte Agrobakterien können dann in bekannter Weise zur Transformation von Pflanzen, insbesondere von Kulturpflanzen verwendet werden, indem beispielsweise verwundete Blätter oder Blattstücke in einer Agrobakterienlösung gebadet und anschließend in geeigneten Medien kultiviert werden.

Die Transformation von Pflanzen durch Agrobakterien ist unter anderem bekannt aus F.F. White, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von S.D. Kung und R. Wu, Academic Press, 1993, S. 15-38. Aus den transformierten Zellen der verwundeten Blätter bzw. Blattstücke können in bekannter Weise transgene Pflanzen regeneriert werden, die ein oder mehrere in die Expressionskassette integrierte Gene enthalten.

Zur Transformation einer Wirtspflanze mit einem oder mehreren erfindungsgemäßen Effektgenen wird eine Expressionskassette als Insertion in einen rekombinanten Vektor eingebaut, dessen Vektor-DNA zusätzliche funktionelle Regulationssignale, beispielsweise Sequenzen für Replikation oder Integration enthält. Geeignete Vektoren sind unter anderem in "Methods in Plant Molecular Biology and Biotechnology" (CRC Press), Kap. 6/7, S. 71-119 (1993) beschrieben.

Unter Verwendung der oben zitierten Rekombinations- und Klonierungstechniken können die Expressionskassetten in geeignete Vektoren kloniert werden, die ihre Vermehrung, beispielsweise in *E. coli*, ermöglichen. Geeignete Klonierungsvektoren sind u.a. pJIT117 (Guerineau et al. (1988) Nucl. Acids Res.16:11380), pBR332, pUC-Serien,

50

M13mp-Serien und pACYC184. Besonders geeignet sind binäre Vektoren, die sowohl in E. coli als auch in Agrobakterien replizieren können.

Die Erfindung betrifft daher weiterhin eine genetisch veränderte Pflanze der Gattung Tagetes, enthaltend einen erfindungsgemäßen Promotor und funktionell verknüpft ein 5 zu eprimierendes Gen, mit der Maßgabe, dass Gene aus Pflanzen der Gattung Tagetes, die in Wildtyppflanzen der Gattung Tagetes von dem jeweiligen Promotor exprimiert werden, ausgenommen sind.

Bevorzugte erfindungsgemäße Promotoren und bevorzugte Effektgene sind vorste-10 hend beschrieben.

Insbesondere bevorzugt sind Effektgene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine β-Hydroxylase, Nukleinsäuren kodierend eine β-Cyclase, Nukleinsäuren kodierend eine ε-Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-

Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat-Δ-20 Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine 25 Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nuk-

leinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.

Bevorzugte, genetisch veränderte Pflanzen der Gattung Tagetes sind Marigold, Tagetes erecta, Tagetes patula, Tagetes lucida, Tagetes pringlei, Tagetes palmeri, Tagetes minuta oder Tagetes campanulata.

Durch die erfindungsgemäßen Promotoren ist es mit Hilfe der vorstehend beschriebenen, erfindungsgemäßen Verfahren möglich, in den vorstehend beschriebenen, erfindungsgemäßen genetisch veränderten Pflanzen der Gattung Tagetes die Stoffwechselwege zu spezifischen biosynthetischen Produkten zu regulieren.

Dazu werden beispielsweise Stoffwechselwege, die zu einem spezifischen biosynthetischen Produkt führen, durch Verursachung oder Erhöhung der Transkriptionrate bzw.

40 Expressionsrate von Genen dieses Biosyntheseweges verstärkt, indem die erhöhte

25

35

51

Proteinmenge zu einer erhöhten Gesamtaktivität dieser Proteine des gewünschten Biosyntheseweges und damit durch einem verstärkten Stoffwechselfluß zu dem gewünschen biosynthetischen Produkt führt.

Je nach gewünschtem biosynthetischen Produkt muss die Transkriptionsrate bzw. Expressionsrate verschiedener Gene erhöht bzw. reduziert werden. In der Regel ist es vorteilhaft, die Transkrioptionsrate bzw. Expressionsrate mehrere Gene zu verändern, d.h. die Transkrioptionsrate bzw. Expressionsrate einer Kombination von Gene zu Erhöhen und/oder die Transkrioptionsrate bzw. Expressionsrate einer Kombination von Gene zu reduzieren.

In den erfindungsgemäßen genetisch veränderten Pflanzen ist mindestens eine erhöhte oder verursachte Expressionsrate eines Gens auf einen erfindungsgemäßen Promotor zurückzuführen.

Weitere, zusätzliche veränderte, d.h. zusätzlich erhöhte oder zusätzlich reduzierte Expressionsraten von weiteren Genen in genetisch veränderten Pflanzen können, müssen aber nicht auf die erfindungsgemäßen Promotoren zurück gehen.

Die Erfindung betrifft daher ein Verfahren zur Herstellung von biosynthetischen Produkten durch Kultivierung von erfindungsgemäßen, genetisch veränderten Pflanzen der Gattung Tagetes.

Die Erfindung betrifft insbesondere ein Verfahren zur Herstellung von Carotinoiden durch Kultivierung von erfindungsgemäßen genetisch veränderten Pflanzen der Gattung Tagetes, dadurch gekennzeichnet, dass die zu exprimierenden Gene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine β-Hydroxylase, Nukleinsäuren kodierend eine β-Cyclase, Nukleinsäuren kodierend eine ε-Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat-Δ-Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein

10

5

20

35

40

52

FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.

Die Carotinoide sind vorzugsweise ausgewählt aus der Gruppe Phytoen, Phytofluen, Lycopin, Lutein, Zeaxanthin, Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

Insbeondere betrifft die Erfindung weiterhin ein Verfahren zur Herstellung von Ketocarotinoiden durch Kultivierung von erfindungsgemäßen genetisch veränderten Pflanzen der Gattung Tagetes, dadurch gekennzeichnet, dass die zu exprimierenden Gene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine  $\beta$ -Hydroxylase, Nukleinsäuren kodierend eine  $\beta$ -Cyclase, Nukleinsäuren kodierend eine  $\epsilon$ -Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat- $\Delta$ -Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.

Die Ketocarotinoide sind vorzugsweise ausgewählt aus der Gruppe Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

Im erfindungsgemäßen Verfahren zur Herstellung von biosynthetischen Produkten, insbesondere Carotinoiden, vorzugsweise Ketocarotinoiden, wird vorzugsweise dem Kultivierungsschritt der genetisch veränderten Pflanzen ein Ernten der Pflanzen und ein Isolieren der biosynthetischen Produkte, insbesondere Carotinoide, vorzugsweise Ketocarotinoide aus den Pflanzen, vorzugsweise aus den Petalen der Pflanzen, angeschlossen.

Die genetisch veränderten Pflanzen der Gattung Tagetes werden in an sich bekannter Weise auf Nährböden gezogen und entsprechend geerntet.

Die Isolierung von Ketocarotinoiden aus den geernteten Blütenblättern erfolgt beispielsweise in an sich bekannter Weise, beispielsweise durch Trocknung und an-

schließender Extraktion und gegebenenfalls weiterer chemischer oder physikalischer Reinigungsprozesse, wie beispielsweise Fällungsmethoden, Kristallographie, thermische Trennverfahren, wie Rektifizierverfahren oder physikalische Trennverfahren, wie beispielsweise Chromatographie. Die Isolierung von Ketocarotinoiden aus den Blütenblättern erfolgt beispielsweise bevorzugt durch organische Lösungsmittel wie Aceton, Hexan, Heptan, Ether oder tert.-Methylbutylether.

Weitere Isolierverfahren von Ketocarotinoiden, insbesondere aus Blütenblättern, sind beispielsweise in Egger und Kleinig (Phytochemistry (1967) 6, 437-440) und Egger (Phytochemistry (1965) 4, 609-618) beschrieben.

Ein besonders bevorzugtes Ketocarotinoid ist Astaxanthin.

Die Ketocarotinoide fallen im erfindungsgemäßen Verfahren in Blütenblättern in Form ihrer Mono- oder Diester mit Fettsäuren an. Einige nachgewiesene Fettsäuren sind z.B. Myristinsäure, Palmitinsäure, Stearinsäure, Ölsäure, Linolensäure, und Laurinsäure (Kamata und Simpson (1987) Comp. Biochem. Physiol. Vol. 86B(3), 587-591).

Von Menschen und Tieren verzehrbare erfindungsgemäße, genetisch veränderte Pflanzen oder Pflanzenteile, wie insbesondere Blütenblätter mit erhöhtem Gehalt an biosynthetischen Produkten, insbesondere Carotinoide, insbesondere Ketocarotinoide, insbesondere Astaxanthin, können auch beispielsweise direkt oder nach an sich bekannter Prozessierung als Nahrungsmittel oder Futtermittel oder als Futter– und Nahrungsergänzungsmittel verwendet werden.

Ferner können die genetisch veränderten Pflanzen zur Herstellung von biosynthetischen Produkt-, insbesondere Carotinoid-, insbesondere Ketocarotinoid-, insbesondere Astaxanthin-haltigen Extrakten und/oder zur Herstellung von Futter- und Nahrungsergänzungsmitteln sowie von Kosmetika und Pharmazeutika verwendet werden.

Die genetisch veränderten Pflanzen der Gattung Tagetes weisen im Vergleich zum Wildtyp einen erhöhten Gehalt an dem gewünschten biosynthetischen Produkten, insbesondere Carotinoide, insbesondere Ketocarotinoide, insbesondere Astaxanthin auf.

Unter einem erhöhten Gehalt wird in diesem Fall auch ein verursachter Gehalt an Ketocarotinoiden, bzw. Astaxanthin verstanden.

Die Erfindung wird durch die nun folgenden Beispiele erläutert, ist aber nicht auf diese beschränkt:

15

20

25

5

10

30

Allgemeine Experimentelle Bedingungen: Sequenzanalyse rekombinanter DNA

Die Sequenzierung rekombinanter DNA-Moleküle erfolgte mit einem Laserfluoreszenz-DNA-Sequenzierer der Firma Licor (Vertrieb durch MWG Biotech, Ebersbach) nach der Methode von Sanger (Sanger et al., Proc. Natl. Acad. Sci. USA 74 (1977), 5463-5467).

### Beispiel 1:

5

10

20

25

35

Amplifikation einer DNA, die die gesamte Primärsequenz der NOST-Ketolase aus Nostoc sp. PCC 7120 kodiert

Die DNA, die für die NOST-Ketolase aus *Nostoc sp. PCC 7120* kodiert, wurde mittels PCR aus *Nostoc sp. PCC 7120* (Stamm der "Pasteur Culture Collection of Cyanobacterium") amplifiziert.

Für die Präparation von genomischer DNA aus einer Suspensionskultur von *Nostoc sp. PCC 7120*, die 1 Woche mit Dauerlicht und konstantem Schütteln (150 rpm) at 25°C in *BG 11-*Medium (1.5 g/l NaNO3, 0.04 g/l K2PO4x3H2O, 0.075 g/l MgSO4xH2O, 0.036 g/l CaCl2x2H2O, 0.006 g/l citric acid, 0.006 g/l Ferric ammonium citrate, 0.001 g/l ED-TA disodium magnesium, 0.04 g/l Na2CO3, 1ml trace metal mix A5+Co (2.86 g/l H3BO3, 1.81 g/l MnCl2x4H2o, 0.222 g/l ZnSO4x7H2o,0.39 g/l NaMoO4X2H2o, 0.079 g/l CuSO4x5H2O, 0.0494 g/l Co(NO3)2x6H2O)) gewachsen war, wurden die Zellen durch Zentrifugation geerntet, in flüssigem Stickstoff eingefroren und im Mörser pulverisiert.

Protokoll für DNA Isolation aus Nostoc PCC7120:

Aus einer 10 ml Flüssigkultur wurden die Bakterienzellen durch 10minütige Zentrifugation bei 8 000 rpm pelletiert. Anschließend wurden die Bakterienzellen in flüssigem Stickstoff mit einem Mörser zerstoßen und gemahlen. Das Zellmaterial wurde in 1 ml 10mM Tris HCl (pH 7.5) resuspendiert und in ein Eppendorf Reaktionsgefäß (2ml Volumen) überführt. Nach Zugabe von 100 µl Proteinase K (Konzentration: 20 mg/ml) wurde die Zellsuspension für 3 Stunden bei 37°C inkubiert. Anschließend wurde die Suspension mit 500 µl Phenol extrahiert. Nach 5minütiger Zentrifugation bei 13 000 upm wurde die obere, wässrige Phase in ein neues 2 ml-Eppendorf Reaktionsgefäß überführt. Die Extraktion mit Phenol wurde 3mal wiederholt. Die DNA wurde durch Zugabe von 1/10 Volumen 3 M Natriumacetat (pH 5.2) und 0.6 Volumen Isopropanol gefällt und anschließend mit 70% Ethanol gewaschen. Das DNA-Pellet wurde bei Raumtemperatur getrocknet, in 25 µl Wasser aufgenommen und unter Erhitzung auf 65°C

gelöst.

5

Die Nukleinsäure, kodierend eine Ketolase aus Nostoc PCC 7120, wurde mittels "polymerase chain reaction" (PCR) aus Nostoc sp. PCC 7120 unter Verwendung eines sense-spezifischen Primers (NOSTF, SEQ ID No. 79) und eines antisensespezifischen Primers (NOSTG SEQ ID No. 80) amplifiziert.

Die PCR-Bedingungen waren die folgenden:

INCIDE UNIDITI & CU. NUBA

- Die PCR zur Amplifikation der DNA, die für ein Ketolase Protein bestehend aus der 10 gesamten Primärsequenz kodiert, erfolgte in einem 50 ul Reaktionsansatz, in dem enthalten war:
  - 1 ul einer Nostoc sp. PCC 7120 DNA (hergestellt wie oben beschrieben)
  - 0.25 mM dNTPs
  - 0.2 mM NOSTF (SEQ ID No. 79)
  - 0.2 mM NOSTG (SEQ ID No. 80)
  - 5 ul 10X PCR-Puffer (TAKARA)
  - 0.25 ul R Taq Polymerase (TAKARA)
- 20 - 25.8 ul Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X94°C 2 Minuten

25 35X94°C 1 Minute

55°C 1 Minuten

72°C 3 Minuten

1X72°C 10 Minuten

Die PCR-Amplifikation mit SEQ ID No. 79 und SEQ ID No. 80 resultierte in einem 805 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (SEQ ID No. 81). Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pGEM-T (Promega) kloniert und der Klon pNOSTF-G erhalten.

35

40

Sequenzierung des Klons pNOSTF-G mit dem M13F- und dem M13R-Primer bestätigte eine Sequenz, welche mit der DNA-Sequenz von 88,886-89,662 des Datenbankeintrages AP003592 identisch ist. Diese Nukleotidsequenz wurde in einem unabhängigem Amplifikationsexperiment reproduziert und repräsentiert somit die Nukleotidsequenz im verwendeten Nostoc sp. PCC 7120.

Beispiel 2

Amplifikation einer DNA, die die gesamte Primärsequenz der NP196-Ketolase aus Nostoc *punctiforme ATCC 29133* kodiert

Die DNA, die für die NP196-Ketolase aus *Nostoc punctiforme ATCC 29133* kodiert, wurde mittels PCR aus *Nostoc punctiforme ATCC 29133* (Stamm der "American Type Culture Collection") amplifiziert.

10

5

Für die Präparation von genomischer DNA aus einer Suspensionskultur von *Nostoc punctiforme ATCC 29133*, die 1 Woche mit Dauerlicht und konstantem Schütteln (150 rpm) at 25°C in *BG 11-*Medium (1.5 g/l NaNO<sub>3</sub>, 0.04 g/l  $K_2$ PO<sub>4</sub>x3H<sub>2</sub>O, 0.075 g/l MgSO<sub>4</sub>xH<sub>2</sub>O, 0.036 g/l CaCl<sub>2</sub>x2H<sub>2</sub>O, 0.006 g/l citric acid, 0.006 g/l Ferric ammonium citrate, 0.001 g/l EDTA disodium magnesium, 0.04 g/l Na<sub>2</sub>CO<sub>3</sub>, 1ml Trace Metal Mix "A5+Co" (2.86 g/l H<sub>3</sub>BO<sub>3</sub>, 1.81 g/l MnCl<sub>2</sub>x4H<sub>2</sub>O, 0.222 g/l ZnSO<sub>4</sub>x7H<sub>2</sub>O, 0.39 g/l Na-MoO<sub>4</sub>X2H<sub>2</sub>O, 0.079 g/l CuSO<sub>4</sub>x5H<sub>2</sub>O, 0.0494 g/l Co(NO<sub>3</sub>)<sub>2</sub>x6H<sub>2</sub>O)) gewachsen war, wurden die Zellen durch Zentrifugation geerntet, in flüssigem Stickstoff eingefroren und im Mörser pulverisiert.

20

25

Protokoll für die DNA-Isolation aus Nostoc punctiforme ATCC 29133:

Aus einer 10 ml Flüssigkultur wurden die Bakterienzellen durch 10 minütige Zentrifugation bei 8000 rpm pelletiert. Anschließend wurden die Bakterienzellen in flüssigem Stickstoff mit einem Mörser zerstoßen und gemahlen. Das Zellmaterial wurde in 1 ml 10mM Tris-HCl (pH 7.5) resuspendiert und in ein Eppendorf-Reaktionsgefäß (2ml Volumen) überführt. Nach Zugabe von 100 µl Proteinase K (Konzentration: 20 mg/ml) wurde die Zellsuspension für 3 Stunden bei 37°C inkubiert. Anschließend wurde die Suspension mit 500 µl Phenol extrahiert. Nach 5minütiger Zentrifugation bei 13 000 upm wurde die obere, wässrige Phase in ein neues 2 ml-Eppendorf-Reaktionsgefäß überführt. Die Extraktion mit Phenol wurde 3mal wiederholt. Die DNA wurde durch Zugabe von 1/10 Volumen 3 M Natriumacetat (pH 5.2) und 0.6 Volumen Isopropanol gefällt und anschließend mit 70% Ethanol gewaschen. Das DNA-Pellet wurde bei Raumtemperatur getrocknet, in 25 µl Wasser aufgenommen und unter Erhitzung auf 65°C gelöst.

35

Die Nukleinsäure, kodierend eine Ketolase aus *Nostoc punctiforme ATCC 29133*, wurde mittels "polymerase chain reaction" (PCR) aus *Nostoc punctiforme ATCC 29133* unter Verwendung eines sense-spezifischen Primers (NP196-1, SEQ ID No. 82) und

eines antisense-spezifischen Primers (NP196-2 SEQ ID No. 83) amplifiziert.

Die PCR-Bedingungen waren die folgenden:

- 5 Die PCR zur Amplifikation der DNA, die für ein Ketolase Protein bestehend aus der gesamten Primärsequenz kodiert, erfolgte in einem 50 ul Reaktionsansatz, in dem enthalten war:
  - 1 ul einer Nostoc punctiforme ATCC 29133 DNA (hergestellt wie oben beschrieben)
- 10 0.25 mM dNTPs
  - 0.2 mM NP196-1 (SEQ ID No. 82)
  - 0.2 mM NP196-2 (SEQ ID No. 83)
  - 5 ul 10X PCR-Puffer (TAKARA)
  - 0.25 ul R Taq Polymerase (TAKARA)
  - 25.8 ul Ag. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X94°C 2 Minuten

20 35X 94°C 1 Minute

35

55°C 1 Minuten

72°C 3 Minuten

1X72°C 10 Minuten

Die PCR-Amplifikation mit SEQ ID No. 82 und SEQ ID No. 83 resultierte in einem 792 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (NP196, SEQ ID No. 84). Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und der Klon pNP196 erhalten.

Sequenzierung des Klons pNP196 mit dem M13F- und dem M13R-Primer bestätigte eine Sequenz, welche mit der DNA-Sequenz von 140.571-139.810 des Datenbank- eintrages NZ\_AABC01000196 identisch ist (inverse orientiert zum veröffentlichen Datenbankeintrag) mit der Ausnahme, daß G in Position 140.571 durch A ersetzt wurde, um ein Standard-Startkodon ATG zu erzeugen. Diese Nukleotidsequenz wurde in einem unabhängigem Amplifikationsexperiment reproduziert und repräsentiert somit die Nukleotidsequenz im verwendeten *Nostoc punctiforme ATCC 29133.* 

Dieser Klon pNP196 wurde daher für die Klonierung in den Expressionsvektor pJIT117(Guerineau et al. 1988, Nucl. Acids Res. 16: 11380) verwendet.

pJIT117 wurde modifiziert, indem der 35S-Terminator durch den OCS-Terminator (Octopine Synthase) des Ti-Plasmides pTi15955 von Agrobacterium tumefaciens (Datenbankeintrag X00493 von Position 12,541-12,350, Gielen et al. (1984) EMBO J. 3 835-846) ersetzt wurde.

Das DNA-Fragment, das die OCS-Terminatorregion beinhaltet, wurde mittels PCR unter Verwendung des Plasmides pHELLSGATE (Datenbankeintrag AJ311874, Wesley et al. (2001) Plant J. 27 581-590, nach Standardmethoden aus *E.coli* isoliert) sowie der Primer OCS-1 (SEQ ID No. 85) und OCS-2 (SEQ ID No. 86) hergestellt.

Die PCR-Bedingungen waren die folgenden:

15

5

10

Die PCR zur Amplifikation der DNA, die die Octopin Synthase (OCS) Terminatorregion (SEQ ID No. 87) beinhaltet, erfolgte in einem 50 ul Reaktionsansatz, in dem enthalten waren:

- 100 ng pHELLSGATE plasmid DNA
- 0.25 mM dNTPs
- 20 0.2 mM OCS-1 (SEQ ID No. 85)
  - 0.2 mM OCS-2 (SEQ ID No. 86)
  - 5 ul 10X PCR-Puffer (Stratagene)
  - 0.25 ul Pfu Polymerase (Stratagene)
  - 28.8 ul Aq. Dest.

25

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:



1X94°C 2 Minuten 35X94°C 1 Minute

50°C 1 Minute

72°C 1 Minute

1X72°C 10 Minuten

Das 210 bp Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und das Plasmid pOCS erhalten.

Sequenzierung des Klons pOCS bestätigte eine Sequenz, die mit einem Sequenzabschnitt auf dem Ti-Plasmid pTi15955 von Agrobacterium tumefaciens (Datenbankeintrag X00493) von Position 12.541 bis 12.350 übereinstimmt.

Die Klonierung erfolgte durch Isolierung des 210 bp Sall-Xhol Fragmentes aus pOCS und Ligierung in den Sall-Xhol geschnittenen Vektor pJIT117.

Dieser Klon heisst pJO und wurde daher für die Klonierung in den Expressionsvektor pJONP196 verwendet.

5

Die Klonierung erfolgte durch Isolierung des 782 Bp Sphl-Fragmentes aus pNP196 und Ligierung in den Sphl geschnittenen Vektor pJO. Der Klon, der die NP196-Ketolase von *Nostoc punctiforme* in der korrekten Orientierung als N-terminale translationale Fusion mit dem rbcS Transitpeptid enthält, heisst pJONP196.

10

#### Beispiel 3:

Herstellung von Expressionsvektoren zur blütenspezifischen Expression der NP196-Ketolase aus Nostoc *punctiforme ATCC 29133* in *Tagetes erecta* 

**1**5

Die Expression der NP196-Ketolase aus Nostoc *punctiforme* in Tagetes erecta erfolgte mit dem Transitpeptid rbcS aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715). Die Expression erfolgte unter Kontrolle des blütenspezifischen Promoters EPSPS aus Petunia hybrida (Datenbankeintrag M37029: Nukleotidregion 7-1787; Benfey et al. (1990) Plant Cell 2: 849-856).

20

Das DNA Fragment, das die EPSPS Promoterregion (SEQ ID No. 88) aus Petunia hybrida beinhaltet, wurde mittels PCR unter Verwendung genomischer DNA (nach Standardmethoden aus Petunia hybrida isoliert) sowie der Primer EPSPS-1 (SEQ ID No. 89) und EPSPS-2 (SEQ ID No. 90) hergestellt.

25

Die PCR-Bedingungen waren die folgenden:



Die PCR zur Amplifikation der DNA, die das EPSPS-Promoterfragment (Datenbankeintrag M37029: Nukleotidregion 7-1787) beinhaltet, erfolgte in einem 50 μl Reaktionsansatz, in dem enthalten war:

- 100 ng genomischer DNA aus A.thaliana
- 0.25 mM dNTPs
- 0.2 mM EPSPS-1 (SEQ ID No. 89)
- 35 0.2 mM EPSPS-2 (SEQ ID No. 90)
  - 5 ul 10X PCR-Puffer (Stratagene)
  - 0.25 ul Pfu Polymerase (Stratagene)
  - 28.8 ul Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X94°C 2 Minuten

35X94°C 1 Minute

5 50°C 1 Minute

5

20

25

35

72°C 2 Minute

1X72°C 10 Minuten

Das 1773 Bp Amplifikat wurde unter Verwendung von Standardmethoden in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und das Plasmid pEPSPS erhalten.

Sequenzierung des Klons pEPSPS bestätigte eine Sequenz, die sich lediglich durch zwei Deletion (Basen ctaagtttcagga in Position 46-58 der Sequenz M37029; Basen aaaaatat in Position 1422-1429 der Sequenz M37029) und die Basenaustausche (T statt G in Position 1447 der Sequenz M37029; A statt C in Position 1525 der Sequenz M37029; A statt G in Position 1627 der Sequenz M37029) von der publizierten EPSPS-Sequenz (Datenbankeintrag M37029: Nukleotidregion 7-1787) unterscheidet. Die zwei Deletionen and die zwei Basenaustausche an den Positionen 1447 und 1627 der Sequenz M37029 wurden in einem unabhängigen Amplifikationsexperiment reproduziert und repräsentieren somit die tatsächliche Nukleotidsequenz in den verwendeten Petunia hybrida Pflanzen.

Der Klon pEPSPS wurde daher für die Klonierung in den Expressionsvektor pJONP196 verwendet.

Die Klonierung erfolgte durch Isolierung des 1763 Bp SacI-HindIII Fragmentes aus pEPSPS und Ligierung in den SacI-HindIII geschnittenen Vektor pJ0NP196. Der Klon, der den Promoter EPSPS anstelle des ursprünglichen Promoters d35S enthält, heisst pJ0ESP:NP196. Diese Expressionskassette enthält das Fragment NP196 in der korrekten Orientierung als N-terminale Fusion mit dem rbcS-Transitpeptid.

Zur Herstellung des Expressionsvektors MSP107 wurde das 2.961 KB bp Sacl-Xhol Fragment aus pJOESP:NP196 mit dem Sacl-Xhol geschnittenen Vektor pSUN3 ligiert MSP 107 beinhaltet Fragment EPSPS den EPSPS Promoter (1761 bp), Fragment *rbcS TP FRAGMENT* das rbcS Transitpeptid aus Erbse (194 bp), Fragment *NP196 KETO CDS* (761 bp), kodierend für die *Nostoc punctiforme* NP196-Ketolase, Fragment *OCS Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

Die Herstellung einer Expressionsvektors für die Agrobacterium-vermittelte Transformation der EPSPS-kontrollierten NP196-Ketolase aus Nostoc punctiforme in Tagetes erecta erfolgte unter der Verwendung des binären Vektors pSUN5 (WO02/00900).

Zur Herstellung des Expressionsvektors MSP108 wurde das 2.961 KB bp Sacl-Xhol Fragment aus pJOESP:NP196 mit dem Sacl-Xhol geschnittenen Vektor pSUN5 ligiert. MSP 108 beinhaltet Fragment EPSPS den EPSPS Promoter (1761 bp), Fragment rbcS TP FRAGMENT das rbcS Transitpeptid aus Erbse (194 bp), Fragment NP196 KETO CDS (761 bp), kodierend für die Nostoc punctiforme NP196-Ketolase, Fragment OCS Terminator (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

## Beispiel 4:

15

20

25

35

Amplifikation einer DNA, die die gesamte Primärsequenz der NP195-Ketolase aus Nostoc punctiforme ATCC 29133 kodiert

Die DNA, die für die NP195-Ketolase aus *Nostoc punctiforme ATCC 29133* kodiert, wurde mittels PCR aus *Nostoc punctiforme ATCC 29133* (Stamm der "American Type Culture Collection") amplifiziert. Die Präparation von genomischer DNA aus einer Suspensionskultur von *Nostoc punctiforme ATCC 29133* wurde in Beispiel 19 beschrieben.

Die Nukleinsäure, kodierend eine Ketolase aus *Nostoc punctiforme ATCC 29133*, wurde mittels "polymerase chain reaction" (PCR) aus *Nostoc punctiforme ATCC 29133* unter Verwendung eines sense-spezifischen Primers (NP195-1, SEQ ID No. 91) und eines antisense-spezifischen Primers (NP195-2 SEQ ID No. 92) amplifiziert.

Die PCR-Bedingungen waren die folgenden:

Die PCR zur Amplifikation der DNA, die für ein Ketolase Protein bestehend aus der gesamten Primärsequenz kodiert, erfolgte in einem 50 ul Reaktionsansatz, in dem enthalten war:

- 1 ul einer Nostoc punctiforme ATCC 29133 DNA (hergestellt wie oben beschrieben)
- 0.25 mM dNTPs
- 0.2 mM NP195-1 (SEQ ID No. 91)
- 0.2 mM NP195-2 (SEQ ID No. 92)
  - 5 ul 10X PCR-Puffer (TAKARA)
  - 0.25 ul R Taq Polymerase (TAKARA)
  - 25.8 ul Aq. Dest.

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X94°C 2 Minuten

35X94°C 1 Minute

55°C 1 Minuten

5

20

25

40

72°C 3 Minuten

1X72°C 10 Minuten

Die PCR-Amplifikation mit SEQ ID No. 91 und SEQ ID No. 92 resultierte in einem 819
Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (NP195, SEQ ID No. 93). Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und der Klon pNP195 erhalten.

Sequenzierung des Klons pNP195 mit dem M13F- und dem M13R-Primer bestätigte eine Sequenz, welche mit der DNA-Sequenz von 55,604-56,392 des Datenbankeintrages NZ\_AABC010001965 identisch ist, mit der Ausnahme, daß T in Position 55.604 durch A ersetzt wurde, um ein Standard-Startkodon ATG zu erzeugen. Diese Nukleotidsequenz wurde in einem unabhängigem Amplifikationsexperiment reproduziert und repräsentiert somit die Nukleotidsequenz im verwendeten *Nostoc punctiforme ATCC 29133*.

Dieser Klon pNP195 wurde daher für die Klonierung in den Expressionsvektor pJ0 (in Beispiel 6 beschrieben) verwendet. Die Klonierung erfolgte durch Isolierung des 809 Bp Sphl-Fragmentes aus pNP195 und Ligierung in den Sphl geschnittenen Vektor pJO. Der Klon, der die NP195-Ketolase von *Nostoc punctiforme* in der korrekten Orientierung als N-terminale translationale Fusion mit dem rbcS Transitpeptid enthält, heisst pJONP195.

# Beispiel 5:

Amplifikation einer DNA, die die gesamte Primärsequenz der NODK-Ketolase aus Nodularia spumignea NSOR10 kodiert.

Die DNA, die für die Ketolase aus *Nodularia spumignea NSOR10* kodiert, wurde mittels PCR aus *Nodularia spumignea NSOR10* amplifiziert.

Für die Präparation von genomischer DNA aus einer Suspensionskultur von *Nodularia spumignea NSOR10*, die 1 Woche mit Dauerlicht und konstantem Schütteln (150 rpm) at 25°C in *BG 11*-Medium (1.5 g/l NaNO<sub>3</sub>, 0.04 g/l K<sub>2</sub>PO<sub>4</sub>x3H<sub>2</sub>O, 0.075 g/l MgSO<sub>4</sub>xH<sub>2</sub>O, 0.036 g/l CaCl<sub>2</sub>x2H<sub>2</sub>O, 0.006 g/l citric acid, 0.006 g/l Ferric ammonium citrate, 0.001 g/l

EDTA disodium magnesium, 0.04 g/l Na $_2$ CO $_3$ , 1ml Trace Metal Mix "A5+Co" (2.86 g/l H $_3$ BO $_3$ , 1.81 g/l MnCl $_2$ x4H $_2$ o, 0.222 g/l ZnSO $_4$ x7H $_2$ 0, 0.39 g/l NaMoO $_4$ X2H $_2$ O, 0.079 g/l CuSO $_4$ x5H $_2$ O, 0.0494 g/l Co(NO $_3$ ) $_2$ x6H $_2$ O) gewachsen war, wurden die Zellen durch Zentrifugation geemtet, in flüssigem Stickstoff eingefroren und im Mörser pulverisiert.

5

10

15

20

25

Protokoll für die DNA-Isolation aus Nodularia spumignea NSOR10:

Aus einer 10 ml Flüssigkultur wurden die Bakterienzellen durch 10 minütige Zentrifugation bei 8000 rpm pelletiert. Anschließend wurden die Bakterienzellen in flüssigem Stickstoff mit einem Mörser zerstoßen und gemahlen. Das Zellmaterial wurde in 1 ml 10mM Tris\_HCl (pH 7.5) resuspendiert und in ein Eppendorf-Reaktionsgefäß (2ml Volumen) überführt. Nach Zugabe von

100 μl Proteinase K (Konzentration: 20 mg/ml) wurde die Zellsuspension für 3 Stunden bei 37°C inkubiert. Anschließend wurde die Suspension mit 500 μl Phenol extrahiert. Nach 5minütiger Zentrifugation bei 13 000 upm wurde die obere, wässrige Phase in ein neues 2 ml-Eppendorf-Reaktionsgefäß überführt. Die Extraktion mit Phenol wurde 3mal wiederholt. Die DNA wurde durch Zugabe von 1/10 Volumen 3 M Natriumacetat (pH 5.2) und 0.6 Volumen Isopropanol gefällt und anschließend mit 70% Ethanol gewaschen. Das DNA-Pellet wurde bei Raumtemperatur getrocknet, in 25 μl Wasser aufgenommen und unter Erhitzung auf 65°C gelöst.

Die Nukleinsäure, kodierend eine Ketolase aus *Nodularia spumignea NSOR10*, wurde mittels "polymerase chain reaction" (PCR) aus *Nodularia spumignea NSOR10* unter Verwendung eines sense-spezifischen Primers (NODK-1, SEQ ID No. 94) und eines antisense-spezifischen Primers (NODK-2 SEQ ID No. 95) amplifiziert.

Die PCR-Bedingungen waren die folgenden:

30

Die PCR zur Amplifikation der DNA, die für ein Ketolase Protein bestehend aus der gesamten Primärsequenz kodiert, erfolgte in einem 50 ul Reaktionsansatz, in dem enthalten war:

- 1 ul einer Nodularia spumignea NSOR10 DNA (hergestellt wie oben beschrieben)
- 0.25 mM dNTPs
- 0.2 mM NODK-1 (SEQ ID No. 94)
- 0.2 mM NODK-2 (SEQ ID No. 95)
- 5 ul 10X PCR-Puffer (TAKARA)
- 0.25 ul R Taq Polymerase (TAKARA)
- 25.8 ul Aq. Dest.

40

35

Die PCR wurde unter folgenden Zyklusbedingungen durchgeführt:

1X94°C 2 Minuten

35X94°C 1 Minute

5 55°C 1 Minuten

72°C 3 Minuten

1X72°C 10 Minuten

Die PCR-Amplifikation mit SEQ ID No. 94 und SEQ ID No. 95 resultierte in einem 720 Bp-Fragment, das für ein Protein bestehend aus der gesamten Primärsequenz kodiert (NODK, SEQ ID No. 96). Unter Verwendung von Standardmethoden wurde das Amplifikat in den PCR-Klonierungsvektor pCR 2.1 (Invitrogen) kloniert und der Klon pNODK erhalten.

Sequenzierung des Klons pNODK mit dem M13F- und dem M13R-Primer bestätigte eine Sequenz, welche mit der DNA-Sequenz von 2130-2819 des Datenbank-eintrages AY210783 identisch ist (inverse orientiert zum veröffentlichen Datenbankeintrag). Diese Nukleotidsequenz wurde in einem unabhängigem Amplifikationsexperiment reproduziert und repräsentiert somit die Nukleotidsequenz im verwendeten *Nodularia spumignea NSOR10*.

Dieser Klon pNODK wurde daher für die Klonierung in den Expressionsvektor pJ0 (in Beispiel 6 beschrieben) verwendet. Die Klonierung erfolgte durch Isolierung des 710 Bp Sphl-Fragmentes aus pNODK und Ligierung in den Sphl geschnittenen Vektor pJO. Der Klon, der die NODK-Ketolase von *Nodularia spumignea* in der korrekten Orientierung als N-terminale translationale Fusion mit dem rbcS Transitpeptid enthält, heisst pJONODK.

#### Beispiel 6:

Herstellung von Expressionsvektoren zur blütenspezifischen Expression der NODK-Ketolase aus *Nodularia spumignea NSOR10* in *Lycopersicon esculentum* und *Tagetes erecta*.

Die Expression der NODK-Ketolase aus *Nodularia spumignea NSOR10* in L. esculentum und Tagetes erecta erfolgte mit dem Transitpeptid rbcS aus Erbse (Anderson et al. 1986, Biochem J. 240:709-715). Die Expression erfolgte unter Kontrolle des blütenspezifischen Promoters EPSPS aus Petunia hybrida (Datenbankeintrag M37029: Nukleotidregion 7-1787; Benfey et al. (1990) Plant Cell 2: 849-856).

20

25

Der Klon pEPSPS (in Beispiel 8 beschrieben) wurde daher für die Klonierung in den Expressionsvektor pJONODK (in Beispiel 12 beschrieben) verwendet.

Die Klonierung erfolgte durch Isolierung des 1763 Bp SacI-HindIII Fragmentes aus pEPSPS und Ligierung in den SacI-HindIII geschnittenen Vektor pJ0NODK. Der Klon, der den Promoter EPSPS anstelle des ursprünglichen Promoters d35S enthält, heisst pJ0ESP:NODK. Diese Expressionskassette enthält das Fragment NODK in der korrekten Orientierung als N-terminale Fusion mit dem rbcS-Transitpeptid.

Die Herstellung eines Expressionsvektors für die Agrobacterium-vermittelte Transformation der EPSPS-kontrollierten NODK-Ketolase aus Nodularia spumignea NSOR10 in L. esculentum erfolgte unter der Verwendung des binären Vektors pSUN3 (WO02/00900).

Zur Herstellung des Expressionsvektors MSP115 wurde das 2.889 KB bp Sacl-Xhol Fragment aus pJOESP:NODK mit dem Sacl-Xhol geschnittenen Vektor pSUN3 ligiert. MSP 115 beinhaltet Fragment EPSPS den EPSPS Promoter (1761 bp), Fragment *rbcS TP FRAGMENT* das rbcS Transitpeptid aus Erbse (194 bp), Fragment *NODK KETO CDS* (690 bp), kodierend für die *Nodularia spumignea NSOR10* NODK-Ketolase,

20 Fragment OCS Terminator (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

Die Herstellung einer Expressionsvektors für die Agrobacterium-vermittelte Transformation der EPSPS-kontrollierten NODK-Ketolase aus *Nodularia spumignea NSOR10* in Tagetes erecta erfolgte unter der Verwendung des binären Vektors pSUN5 (WO02/00900).

Zur Herstellung des Expressionsvektors MSP116 wurde das 2.889 KB bp SacI-Xhol Fragment aus pJOESP:NODK mit dem SacI-Xhol geschnittenen Vektor pSUN5 ligiert. MSP 116 beinhaltet Fragment EPSPS den EPSPS Promoter (1761 bp), Fragment *rbcS TP FRAGMENT* das rbcS Transitpeptid aus Erbse (194 bp), Fragment *NODK KETO CDS* (690 bp), kodierend für die *Nodularia spumignea NSOR10* NODK-Ketolase, Fragment *OCS Terminator* (192 bp) das Polyadenylierungssignal von Octopin-Synthase.

Beispiel 7:

35

Herstellung transgener Tagetes Pflanzen

Tagetessamen werden sterilisiert und auf Keimungsmedium (MS-Medium; Murashige and Skoog, Physiol. Plant. 15(1962), 473-497) pH 5,8, 2% Saccharose) aufgelegt. Die

10

15

20

25

35

66

Keimung erfolgt in einem Temperatur/Licht/Zeitintervall von 18-28°C/20-200  $\mu$ E/3 - 16 Wochen, bevorzugt jedoch bei 21°C, 20-70  $\mu$ E, für 4-8 Wochen.

Alle Blätter der sich bis dahin entwickelten in vitro Pflanzen werden geerntet und quer zur Mittelrippe geschnitten. Die dadurch entstehenden Blattexplantate mit einer Größe von 10 - 60 mm² werden im Verlaufe der Präparation in flüssigem MS - Medium bei Raumtemperatur für maximal 2 h aufbewahrt.

Ein beliebiger Agrobakterium tumefaciens Stamm, bevorzugt aber ein supervirulenter Stamm, wie z.B. EHA105 mit einem entsprechenden Binärplasmid, das ein Selektionsmarkergen (bevorzugt *bar* oder *pat*) sowie ein oder mehrere Trait- oder Reportergene tragen kann wird (pS5FNR:NOST,pS5AP3:NOST pS5FNR:NP196, pS5EPS:NP196, pS5FNR:NP195, pS5EPS:NP195, pS5FNR:NODK und pS5EPS:NODK), über Nacht angezogen und für die Co-Kultivierung mit dem Blattmaterial verwendet. Die Anzucht des Bakterienstammes kann wie folgt erfolgen: Eine Einzelkolonie des entsprechenden Stammes wird in YEB (0,1 % Hefeextrakt, 0,5 % Rindfleischextrakt, 0,5 % Pepton, 0,5 % Saccharose, 0,5 % Magnesiumsulfat x 7 H<sub>2</sub>0) mit 25 mg/l Kanamycin angeimpft und bei 28°C für 16 bis 20 h angezogen. Anschließend wird die Bakteriensuspension durch Zentrifugation bei 6000 g für 10 min geerntet und derart in flüssigem MS Medium resuspendiert, daß eine OD600 von ca. 0,1 bis 0,8 entstand. Diese Suspension wird fuer die C-Kultivierung mit dem Blattmaterial verwendet.

Unmittelbar vor der Co-Kultivierung wird das MS-Medium, in dem die Blätter aufbewahrt worden sind, durch die Bakteriensuspension ersetzt. Die Inkubation der Blättchen in der Agrobakteriensuspension erfolgte für 30 min unter leichtem Schütteln bei Raumtemperatur. Anschließend werden die infizierten Explantate auf ein mit Agar (z.B. 0,8 % Plant Agar (Duchefa, NL) verfestigtes MS-Medium mit Wachstumsregulatoren, wie beispielsweise 3 mg/l Benzylaminopurin (BAP) sowie 1 mg/l Indolylessigsäure (IAA) aufgelegt. Die Orientierung der Blätter auf dem Medium ist bedeutungslos. Die Kultivierung der Explantate findet für 1 bis 8 Tage, bevorzugt aber für 6 Tage statt, dabei können folgende Bedingungen angewendet werden: Lichtintensität: 30-80 $\mu \text{Mol/m}^2 \text{ x sec}$ , Temperatur: 22 – 24 $^{\circ}\text{C}$ , hell/dunkel Wechsel von 16/8 Stunden. Anschließend werden die co-kultivierten Explantate auf frisches MS-Medium, bevorzugt mit den gleichen Wachstumsregulatoren übertragen, wobei dieses zweite Medium zusätzlich ein Antibiotikum zur Unterdrückung des Bakterienwachstums enthält. Timentin in einer Konzentration von 200 bis 500 mg/l ist für diesen Zweck sehr geeignet. Als zweite selektive Komponente wird eine für die Selektion des Transformationserfolges eingesetzt. Phosphinothricin in einer Konzentration von 1 bis 5 mg/l selektiert sehr effizient, aber auch andere selektive Komponenten gemäß des zu verwendenden Verfah-

20040135

rens sind denkbar.

Nach jeweils ein bis drei Wochen erfolgt der Transfer der Explantate auf frisches Medium bis sich Sproßknospen und kleine Sprosse entwickeln, die dann auf das gleiche Basalmedium einschließlich Timentin und PPT oder alternative Komponenten mit Wachstumsregulatoren, nämlich z.B. 0,5 mg/l Indolylbuttersäure (IBA) und 0,5 mg/l Gibberillinsäure GA<sub>3</sub>, zur Bewurzelung übertragen werden. Bewurzelte Sprosse können ins Gewächshaus überführt werden.

Zusätzlich zu der beschriebenen Methode sind folgende vorteilhafte Modifikationen möglich:

Bevor die Explantate mit den Bakterien infiziert werden, können sie für 1 bis 12 Tage, bevorzugt 3 - 4, auf das oben beschriebene Medium für die Co-Kultur vorinkubiert werden. Anschließend erfolgt die Infektion, Co-Kultur und selektive Regeneration wie oben beschrieben.

Der pH Wert für die Regeneration (normalerweise 5,8) kann auf pH 5,2 gesenkt werden. Dadurch wird die Kontrolle des Agrobakterienwachstums verbessert.

Die Zugabe von AgNO<sub>3</sub> (3 - 10 mg/l) zum Regenerationsmedium verbessert den Zustand der Kultur einschließlich der Regeneration selbst.

Komponenten, die die Phenolbildung reduzieren und dem Fachmann bekannt sind, wie z.B. Zitronensäure, Ascorbinsäure, PVP u.v.a.m., wirken sich positiv auf die Kultur aus.

Für das gesamte Verfahren kann auch flüssiges Kulturmedium Verwendung finden. Die Kultur kann auch auf handelsüblichen Trägern, die auf dem flüssigen Medium positioniert werden inkubiert werden.

Gemäß der oben beschriebenen Transformationsmethode wurden mit folgenden Expressionskonstrukten folgende Linien erhalten:

Mit pS5FNR:NOST wurde beispielsweise erhalten: MSP102-1, MSP102-2, MSP102-3,

Mit pS5AP3:NOST wurde beispielsweise erhalten: MSP104-1, MSP104-2, MSP104-3

Mit pS5FNR:NP196 wurde erhalten: MSP106-1, MSP106-2, MSP106-3

40 Mit pS5EPS:NP196 wurde erhalten: MSP108-1, MSP108-2, MSP108-3

20

5

30

35

Mit pS5FNR:NP195 wurde erhalten: MSP110-1, MSP110-2, MSP110-3

Mit pS5EPS:NP195 wurde erhalten: MSP112-1, MSP112-2, MSP112-3

Mit pS5FNR:NODK wurde erhalten: MSP114-1, MSP114-2, MSP114-3

Mit pS5EPS:NODK wurde erhalten: MSP116-1, MSP116-2, MSP116-3

## 10 Beispiel 8:

5

Enzymatische Lipase-katalysierte Hydrolyse von Carotinoidestern aus Pflanzenmaterial und Identifizierung der Carotinoide

# Allgemeine Arbeitsvorschrift

**1**5

20

25

- a) Gemörsertes Pflanzenmaterial (z.B. Petalenmaterial) (30-100 mg Frischgewicht) wird mit 100% Aceton (dreimal 500µl; jeweils etwa 15 Minuten schütteln) extrahiert. Das Lösungsmittel wird evaporiert. Carotinoide werden anschließend in 495 µl Aceton aufgenommen, 4,95 ml Kaliumphosphatpuffer (100 mM, pH7.4) zugegeben und gut gemischt. Danach erfolgt die Zugabe von ca. 17 mg Bile-Salze (Sigma) und 149 µl einer NaCl/CaCl₂-Lösung (3M NaCl und 75 mM CaCl₂). Die Suspension wird für 30 Minuten bei 37°C inkubiert. Für die enzymatische Hydrolyse der Carotinoidester wird 595 µl einer Lipaselösung (50 mg/ml Lipase Typ7 von Candida rugosa (Sigma)) zugegeben und unter Schütteln bei 37C inkubiert. Nach etwa 21 Stunden erfolgte nochmals eine Zugabe von 595 µl Lipase mit erneuter Inkubation von mindestens 5 Stunden hai azuse.
- Zugabe von 595 μl Lipase mit erneuter Inkubation von mindestens 5 Stunden bei 37°C. Anschließend werden etwa ca. 700 mg Na<sub>2</sub>SO<sub>4</sub> in der Lösung gelöst. Nach Zugabe von 1800 μl Petrolether werden die Carotinoide durch kräftig Mischen in die organische Phase extrahiert. Dieses Ausschütteln wird solange wiederholt, bis die organische Phase farblos bleibt. Die Petroletherfraktionen werden vereinigt und der Petrolether evaporiert. Freie Carotinoide werden in 100-120 μl Aceton aufgenommen. Mittels HPLC und C30-reverse phase-Säule können freie Carotinoide aufgrund von Retentionszeit und UV-VIS-Spektren identifiziert werden.
- Die verwendeten Bile-Salze oder Gallensäuresalze sind 1:1 Mischungen von Cholat und Desoxycholat.
  - b) Arbeitsvorschrift für Aufarbeitung, wenn nur geringe Mengen an Carotinoidestern im Pflanzenmaterial vorhanden sind
- 40 Alternativ kann die Hydrolyse der Carotinoidester durch Lipase aus Candida rugosa nach Trennung mittels Dünnschichtchromatographie erreicht werden. Dazu werden 50-

100mg Pflanzenmaterial dreimal mit etwa 750μl Aceton extrahiert. Der Lösungsmittelextrakt wird im Vakuum einrotiert (erhöhte Temperaturen von 40-50°C sind tolerabel). Danach erfolgt Zugabe von 300μl Petrolether:Aceton (Verhältnis 5:1) und gute Durchmischung. Schwebstoffe werden durch Zentrifugation (1-2 Minuten) sedimentiert. Die obere Phase wird in ein neues Reaktionsgefäß überführt. Das verbleibende Rest wird erneut mit 200μl Petrolether:Aceton (Verhältnis 5:1) extrahiert und Schwebstoffe werden durch Zentrifugation entfernt. Die beiden Extrakte werden zusammengeführt (Volumen 500μl) und die Lösungsmittel evaporiert. Der Rückstand wird in 30μl Petrolether:Aceton (Verhältnis 5:1) resuspendiert und auf eine Dünnschichtplatte (Silica-Gel 60, Merck) aufgetragen. Falls mehr als eine Auftragung für präparativ-analytische Zwecke erforderlich ist, sollten mehrere Aliquots mit jeweils 50-100 mg Frischgewicht in der beschriebenen Weise für die dünnschichtchromatographische Trennung aufbereitet werden.



5

10

20

25

35

40

Die Dünnschichtplatte wird in Petrolether: Aceton (Verhältnis 5:1) entwickelt. Carotinoidbanden können visuell aufgrund ihrer Farbe identifiziert werden. Einzelne Carotinoidbanden werden ausgekratzt und können für präparativ-analytische Zwecke gepoolt werden. Mit Aceton werden die Carotinoide vom Silica-Material eluiert; das Lösungsmittel wird im Vakuum evaporiert. Zur Hydrolyse der Carotinoidester wird der Rückstand in 495µl Aceton gelöst, 17mg Bile-Salze (Sigma), 4,95ml 0.1M Kaliumphosphatpuffer (pH 7,4) und 149µl (3M NaCl, 75mM CaCl<sub>2</sub>) zugegeben. Nach guter Durchmischung wird 30min bei 37°C äquilibriert. Danach erfolgt die Zugabe von 595µl Lipase von Candida rugosa (Sigma, Stammlösung von 50mg/ml in 5mM CaCl<sub>2</sub>). Über Nacht erfolgt die Inkubation mit Lipase unter Schütteln bei 37°C. Nach etwa 21 Stunden wird nochmals die gleiche Menge an Lipase zugegeben; für mindestens 5 Stunden wird nochmals bei 37°C unter Schütteln inkubiert. Dann erfolgt die Zugabe von 700mg Na<sub>2</sub>SO<sub>4</sub> (wasserfrei); mit 1800μl Petrolether wird für ca. 1 Minute ausgeschüttelt und die Mischung bei 3500 Umdrehungen/Minute für 5 Minuten zentrifugiert. Die obere Phase wird in ein neues Reaktionsgefäß überführt und das Ausschütteln so lange wiederholt, bis die obere Phase farblos ist. Die vereinigte Petrolether-Phase wird im Vakuum eingeengt (Temperaturen von 40-50°C sind möglich). Der Rückstand wird in 120µl Aceton, eventuell mittels Ultraschall, gelöst. Die gelösten Carotinoide können mittels HPLC unter Verwendung einer C30-Säule getrennt und anhand von Referenzsubstanzen quantifiziert werden.



Beispiel 9:

HPLC-Analyse freier Carotinoide

Die Analyse der nach der Arbeitsvorschriften in Beispiel 15 erhaltenen Proben erfolgt unter folgenden Bedingungen:

Folgende HPLC-Bedingungen wurden eingestellt.

Trennsäule: Prontosil C30-Säule, 250 x 4,6 mm, (Bischoff, Leonberg, Germany)

Flussrate: 1.0 ml/min

5 Eluenten: Laufmittel A - 100% Methanol

Laufmittel B - 80% Methanol, 0.2% Ammoniumacetat

Laufmittel C - 100% t-Butyl-methylether

Detektion: 300-530 nm

## 10 Gradientenprofil:

Zeit	Flussrate	% Laufmittel A	% Laufmittel B	% Laufmittel C
1.00	1.0	95.0	5.0	· 0
12.00	1.0	95.0	5.0	0
12.10	1.0	80.0	5.0	15.0
22.00	1.0	76.0	5.0	19.0
22.10	1.0	66.5	5.0	28.5
38.00	1.0	15.0	5.0	80.0
45.00	1.0	95.0	5.0	0
46.0	1.0	95.0	5.0	0

Einige typische Retentionszeiten für erfindungsgemäß gebildete Carotinoide sind z.B.: Violaxanthin 11, 7 min, Astaxanthin 17,7 min, Adonixanthin 19 min, Adonirubin 19,9 min, Zeaxanthin 21 min.



### Patentansprüche

- 1. Verwendung eines Promotors, ausgewählt aus der Gruppe
- 5 A) EPSPS Promotor
  - B) B-Gene Promotor
  - C) PDS Promotor und
  - D) CHRC Promotor
- zur Expression von Genen in Pflanzen der Gattung Tagetes, mit der Maßgabe, dass Gene aus Pflanzen der Gattung Tagetes, die in Wildtyppflanzen der Gattung Tagetes von dem jeweiligen Promotor exprimiert werden, ausgenommen sind.



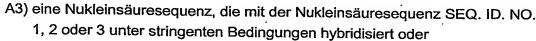
- 2. Verwendung nach Anspruch 1, dadurch gekennzeichnet, dass die Expression spezifisch in Blüten erfolgt.
- 3. Verwendung nach Anspruch 2, dadurch gekennzeichnet, dass die Expression spezifisch in Petalen erfolgt.
- Verwendung nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass der EPSPS Promotor gemäß Ansprüch 1



30

35

- A1) die Nukleinsäuresequenz SEQ. ID. NO. 1, 2 oder 3 oder
- A2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 1, 2 oder 3 aufweist oder



A4) funktionell äquivalente Fragmente der Sequenzen unter A1), A2) oder A3)

enthält.

- 5. Verwendung nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass der B-Gene Promotor gemäß Ansprüch 1
- B1) die Nukleinsäuresequenz SEQ. ID. NO. 4, 5 oder 6 oder 135/2004 Mec... 22.07.2004 Seq



- B2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 4, 5 oder 6 aufweist oder
- B3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO.
  4, 5 oder 6 unter stringenten Bedingungen hybridisiert oder
  B4) funktionell äquivalente Fragmente der Sequenzen unter B1), B2) oder B3)

over B3

enthält.

 Verwendung nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass der PDS Promotor gemäß Anspruch 1

- C1) die Nukleinsäuresequenz SEQ. ID. NO. 7, 8, 9 oder 10 oder
- C2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 7, 8, 9 oder 10 aufweist oder
- C3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO.7, 8, 9 oder 10 unter stringenten Bedingungen hybridisiert oder
- C4) funktionell äquivalente Fragmente der Sequenzen unter C1), C2) oder C3)

enthält.

 Verwendung nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass der CHRC Promotor gemäß Anspruch 1

30

35

5

10

5

20

- D1) die Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13 oder 14 oder
- D2) eine von diesen Sequenzen durch Substitution, Insertion oder Deletion von Nukleotiden abgeleitete Sequenz, die eine Identität von mindestens 60 % auf Nukleinsäureebene mit der jeweiligen Sequenz SEQ. ID. NO. 11, 12, 13 oder 14 aufweist oder
- D3) eine Nukleinsäuresequenz, die mit der Nukleinsäuresequenz SEQ. ID. NO. 11, 12, 13 oder 14 unter stringenten Bedingungen hybridisiert oder
- D4) funktionell äquivalente Fragmente der Sequenzen unter D1), D2) oder D3)

enthält.

10

15

20

25

30

35

- 8. Genetisch veränderte Pflanze der Gattung Tagetes, wobei die genetische Veränderung zu einer Erhöhung oder Verursachung der Expressionsrate mindestens eines Gens im Vergleich zum Wildtyp führt und bedingt ist durch die Regulation der Expression dieses Gens in der Pflanze durch Promotoren gemäß einem der Ansprüche 1 bis 7.
- Genetisch veränderte Pflanze nach Anspruch 8, dadurch gekennzeichnet, dass die Regulation der Expression von Genen in der Pflanze durch Promotoren gemäß einem der Ansprüche 1 bis 7 dadurch erreicht wird, dass man
  - eine oder mehrere Promotoren gemäß einem der Ansprüche 1 bis 7 in das Genom der Pflanze einbringt, so dass die Expression eines oder mehrerer endogenen Gene unter der Kontrolle der eingebrachten Promotoren gemäß einem der Ansprüche 1 bis 7 erfolgt oder
  - ein oder mehrere Gene in das Genom der Pflanze einbringt, so dass die Expression eines oder mehrerer der eingebrachten Gene unter der Kontrolle der endogenen Promotoren gemäß einem der Ansprüche 1 bis 7 erfolgt oder
  - ein oder mehrere Nukleinsäurekonstrukte, enthaltend mindestens einen Promotor gemäß einem der Ansprüche 1 bis 7 und funktionell verknüpft eine oder mehrere, zu exprimierende Gene, in die Pflanze einbringt.
- 10. Genetisch veränderte Pflanze der Gattung Tagetes, enthaltend einen Promotor gemäß einem der Ansprüche 1 bis 7 und funktionell verknüpft ein zu eprimierendes Gen, mit der Maßgabe, dass Gene aus Pflanzen der Gattung Tagetes, die in Wildtyppflanzen der Gattung Tagetes von dem jeweiligen Promotor exprimiert werden, ausgenommen sind.
- 11. Genetisch veränderte Pflanze nach einem der Ansprüche 8 bis 10, dadurch gekennzeichnet, dass die zu exprimierenden Gene ausgewählt sind aus der Gruppe Nukleinsäuren kodierend ein Protein aus den Biosyntheseweg von proteinogenen und nicht-proteinogenen Aminosäuren, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Nukleotiden und Nukleosiden, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von organischen Säuren, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Lipiden und Fettsäuren, Nuk-

leinsäuren kodierend ein Protein aus dem Biosyntheseweg von Diolen, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Konhlenhydraten, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von aromatischen Verbindung, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Vitaminen, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Carotinoiden, Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Cofaktoren und Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Enzymen, wobei die Gene gegebenenfalls weitere Regulationselemente enthalten können.

10

5

- 12. Genetisch veränderte Pflanze nach Anspruch 11, dadurch gekennzeichnet, dass man als zu exprimierende Gene Nukleinsäuren kodierend ein Protein aus dem Biosyntheseweg von Carotinoiden verwendet.
- 13. Genetisch veränderte Pflanze nach Anspruch 12, dadurch gekennzeichnet, dass die zu exprimierenden Gene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine β-Hydroxylase, Nukleinsäuren kodierend eine ε-Cyclase, Nukleinsäuren kodierend eine ε-Cyclase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat-Δ-Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Physiene Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Physiene
  - - 14. Verfahren zur Herstellung von biosynthetischen Produkten durch Kultivierung von genetisch veränderten Pflanzen der Gattung Tagetes gemäß einem der Ansprüche 8 bis 13.

ein FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.

toen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend

10

15

30

- 15. Verfahren zur Herstellung von Carotinoiden durch Kultivierung von genetisch veränderten Pflanzen gemäß einem der Ansprüche 8 bis 13, dadurch gekennzeichnet, dass die zu exprimierenden Gene ausgewählt sind aus der Gruppe Nukleinsäuren, kodierend eine Ketolase, Nukleinsäuren kodierend eine β-Hydroxylase, Nukleinsäuren kodierend eine  $\beta$ -Cyclase, Nukleinsäuren kodierend eine  $\epsilon$ -Cyclase, Nukleinsäuren kodierend eine Epoxidase, Nukleinsäuren kodierend eine HMG-CoA-Reduktase, Nukleinsäuren kodierend eine (E)-4-Hydroxy-3-Methylbut-2-enyl-Diphosphat-Reduktase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Synthase, Nukleinsäuren kodierend eine 1-Deoxy-D-Xylose-5-Phosphat-Reduktoisomerase, Nukleinsäuren kodierend eine Isopentenyl-Diphosphat-Δ-Isomerase, Nukleinsäuren kodierend eine Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Farnesyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Geranyl-Geranyl-Diphosphat-Synthase, Nukleinsäuren kodierend eine Phytoen-Synthase, Nukleinsäuren kodierend eine Phytoen-Desaturase, Nukleinsäuren kodierend eine Prephytoen-Synthase, Nukleinsäuren kodierend eine Zeta-Carotin-Desaturase, Nukleinsäuren kodierend ein crtISO Protein, Nukleinsäuren kodierend ein FtsZ Protein und Nukleinsäuren kodierend ein MinD Protein.
- Verfahren nach Anspruch 15, dadurch gekennzeichnet, dass man nach dem Kulti vieren die genetisch veränderten Pflanzen erntet und anschließend die Carotinoide aus den genetisch veränderten Pflanzen isoliert.
  - 17. Verfahren nach Anspruch 16, dadurch gekennzeichnet, dass man die Blüten der genetisch veränderten Pflanzen erntet und anschließend die Carotinoide aus den Petalen der genetisch veränderten Pflanzen isoliert.
  - 18. Verfahren nach einem der Ansprüche 15 bis 17, dadurch gekennzeichnet, dass die Carotinoide ausgewählt sind aus der Gruppe Phytoen, Lycopin, Lutein, Zeaxanthin, Astaxanthin, Canthaxanthin, Echinenon, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Adonirubin, Violaxanthin und Adonixanthin.

# 04sequ.txt SEQUENCE LISTING

<110>	SunGene GmbH & Co. KGaA	
<120>	Promotoren zur Expression von Genen in Tagetes	
<130>	PF 55341	
<160>	97	
<170>	PatentIn version 3.1	
0>	1	
<211>	174	
<212>	DNA	
<213>	Petunia hybrida	
<220>		
<221>	promoter	
<222>	(1)(174)	
<223>	EPSPS Promotor	
	•	
<400>	1 cacag ctggaatttt ttacaaaggt agttggtgaa gctagtcagc gaatcccatt	60
	ccact ctacctaacc cccttcacca acaacaaatt tctgtaattt aaaaactagc	120
	aagaa ctctcttta caaagagcca aagactcaat ctttactttc aaga	174
<210>		
	1781	
<212>		
<213>	Petunia hybrida	•
<220>		
<221>		
<<<1>><<1>> 1	nramatar	

<400> 2 tagcgcaatc 1	ttatotoota	caaatettea	ttaatcaaac	****		
						60
atgtggccct a						120
gttgaagcat 1						180
cttttatgac a						240
ttttccttcc a						300
catattctct c	Licgittaca	cgattatagt	aataatgata	taggatgaca	gaagttgaca	360
cataaatttt t						420
taatatgtat t						480
gacacctact a						540
				ttgacagtgt		600
-gttaactaa t						660
caataaaaaa t	agaagaatc	ttcctgaggc	aaagttttgg	aaaaattaag	agtggctgag	720
atttaatttc a	acaggaatt	agttccactt	aacttttagg	ttacgataca	gtgctaatta	780
aataacttaa t						840
agcaatcaaa a						900
ctcactgttt g	ttggttaaa	acgtagactt	acacctacca	aaatctacaa	ctaaaatgag	960
gcaataatac t	ttgcccaaa	attaccaaga	aaagaaaaag	aaaggaatcc	cttaatatta	1020
ctctcctcca t	ttcacaata	aatatcctag	tttgacttaa	attagagttt	aaaaaatgaa	1080
agacgacttt t	aaaacttgt	aatctaaaat	aaatcatagt	taaatgtgtg	gctataaatc	1140
attgtattaa c	ggtaaagtg	gtaagtttaa	aagttaattg	ttttcaaata	taaaattota	1200
				catccattat		1260
				aattatcagt		1320
taaaaattat t						1380
gcttttccag g						1440
gtccagactt c						1500
ttcgaattag to						1560
caaaaatgtc aa	aattaatat a	atcaatctgc	aacaaccttt	tcaccttgag	22626265	
gaattttta ca	aaaggtagt 1	tggtgaagct	agtcagcgaa	tcccattacc	ttccactay	1620
cctaaccccc to	tcaccaaca a	acaaatttct	ataatttaan	22C+2666-	- LCCACTCTA	1680
tcttttacaa ag					aaaagaactc	1740
	, .j aaag '		cacciccaag	a		1781

<211> 1760
<212> DNA
<213> Petunia hybrida

<220>
<221> promoter
<222> (1)..(1760)
<223> EPSPS Promotor

<400> 3 tagcgcaatc ttatgtggta caaatcttga ttagtcggga aaaaatgatg tggccctaca 60 aatggttgga ggatgggaga tttggctcta tctagagtta tgtggttgtt gaagcatttg 120 actetet getgtggtag ttggcatate cacattgtet cettecaett ttatgacaat 180 acgtgaaag ttatgggttg ttttgtctat ttttgtcgag gcctttcttt tccttccagg 240 ttgttgaaga tggtccaatt cgattagaat aatgttttga gctttagcat attctctctc 300 gtttacacga ttatagtaat aatgatatag gatgacagaa gttgacacat aaatttttta 360 ttctctccat ttactttaat ccaaatctca cctaccctaa acttctttaa tatgtattca 420 atagtctatc cgagtaaatt gtaaatttaa caaccattga taatattgac acctactaac 480 atatactagt aaagagaata ttaacatggc acatataatt tgatgcaaaa tgagtatgat 540 gaaatttaaa cccaaaatct cttgattttg acagtgtcac cttgacttgt taactaataa 600 gtcatgtttt agtggcagaa agacaaactc atccaccaac tgtatagcaa taaaaaatag 660 aagaatcttc Ctgaggcaaa gttttggaaa aattaagagt gqctgagatt taatttcaac 720 aggaattagt tccacttaac ttttaggtta cgatacagtg ctaattaaat aacttaattg 780 tagatat ttcttgcacc taaaaaattt aaaaactgaa aaaaggtagc aatcaaaata 840 aaaagga caaaataagt gaaaggtaca gccaccaacc ctggcggctc actgtttgtt 900 ggttaaaacg tagacttaca cctaccaaaa tctacaacta aaatgaggca ataatacttt 960 gcccaaaatt accaagaaaa gaaaaagaaa ggaatccctt aatattactc tcctccattt 1020 cacaataaat atcctagttt gacttaaatt agagtttaaa aaatgaaaga cgacttttaa 1080 aacttgtaat ctaaaataaa tcatagttaa atgtgtggct ataaatcatt gtattaacgg 1140 taaagtggta agtttaaaag ttaattgttt tcaaatataa aattgtacta tcattctttt 1200 tggaatggac taataagaaa actatgacat ccattatgga gcggagggag tatctccttt 1260 taacaataac ctttgtccct tcaattcaat tatcagtatg caaacattaa aaattattat 1320 tgatgttaag taccacatca tccttaatga tagaatcatc gtagaacgct tttccaggca 1380 cacattcaaa ctagttagac cagtaccaca catcgaatat tccagacttc tttgtttgaa 1440 tagtcgacta cattggataa tggaacttct cgaattaact tcgaattagt cgagcccaaa 1500

Seite 3

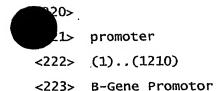
ataatatata	cgtcgggtgg	aaaactataa	aatgtttgac	aaaaatgtca	aattaatata	1560
tcaatctgca	acaacctttt	caccttgaga	acacagctga	aattttttac	aaaggtagtt	1620
ggtgaagcta	gtcagcgaat	cccattacct	tccactctac	ctaaccccct	tcaccaacaa	1680
caaatttctg	taatttaaaa	actagccaaa	aaagaactct	cttttacaaa	gagccaaaga	1740
ctcaatcttt	actttcaaga					1760

<210> 4

<211> 1210

<212> DNA

<213> Lycopersicon esculentum



<400> gaattctctg aaaaggagca ccatatttgc cgcactgtgg ttcatatttc caagtacatt 60 120 tagatgaact atatcatcag attgaaaggt tattgtataa tcaatccagt ggattctcgt 180 tctggcacct ttagaagtac atgtgcggaa aagaatgata aggtttgtat tgttgttgac 240 aaagcctgtt gcctttctca tttgtaaatg ttctgaacga ctcctaaatt actcttaagg 300 tgtaaggtct tccgtgcctg tttgtaaata taatgctgtg ccgtgactta ccttttgtac 360 catttgttca aatgtatggc ctgaacacca gggttgtcaa aaatgtctca tgcccgtttt 420 ggtctga aaatggcgtg atgccaaatt ctgccgctcc acagtgagca tttcgatcta 480 qaaattq accaacttat tttatcactt gataactaaa caaaatccta ttaactttaa tcatacattg tatttatacc gaaaaattta tgcataactc attaaattac cttttttagc 540 600 agtcaaattc taaatcagtt tctaatttat caaaatggct tttatagggt cccatttcca ctaatatacc tgccgtccat gcactgacta caaaacaaat acctcactat gtttgttagt 660 720 gcttggtaat ataaaacctt ttcttttatg agaaagttca ccgagaataa ttttctattt 780 gtggcataat agtatatagt gcagattgac aagaatttaa ttttgcagtt gggcacatga acaattttcc tcaaagttgt agaaagtact tttcattttc ttgtcaccga aaattattta 840 900 taattqaaat taaaaccgaa tgagctgcaa gattcaagtc gaattttcaa aagaattgac 960 caagaaaaaa ttcaaaaata tcccccaccc cctaccaaac acatcctaaa gtgaggtata 1020 qactqqqact qqgattggga aaagggtaaa atgctttcac tagcttagca aagattccac 1080 tttgttagct atctttcttt ctcatttcct tttttctttt tcttttttt gttatataag Seite 4

	04Sequ.txt		
ccaaag	agtagg tacccaaaag catcaatatt ttgtattgct tggtgattc	tctgtagtcc	1140
agtatt	ttcat tttctacaag ttccacctcc ctccataatt aaccattate	aatcttatac	1200
attctc	ctata		1210
<210>	,		1210
<211>	1599		
<212>	DNA		
<213>	Lycopersicon esculentum		
<220>			
<221>	promoter		
922	(1) (4100)		

22> (1)..(1599) 3> B-Gene Promotor

				•			
	<400> 5 atctcattgt	ataocttoto		at cat - Lt.			
	Ctatttcata	Cttctata-a		gregrettag	gcttgggtta	gttggtgttg	60
	2 stanting	. CLICLALCAA	ccttgtgtga	gttcctttat	aaaatatgac	tgttggagga	120
	agtaatttac	ctttagttcg	actacatcaa	gatttgcatc	attctcgtcc	aagaaatctt	180
	agtttgaagc	cttttggtct	ggtatatttg	tcaatctgag	cttcgcaact	ttctcatgac	240
	aggggtttgt	tgacatgcct	gattgtgctc	ttcctttact	tgataattgc	tgcttgttgc	300
	ggaggcatca	ctctaccttc	ctgcagatca	tgaattctct	gaaaaggagc	accatattta	360
	ccgcactgtg	gttcatattt	ccaattacat	ttagatgaac	tatatcatca	anantassa	420
	gttattgtat	aatcaatcca	gtggattctc	gttctggcac	Ctttagaagt	352tataaa	
1	agaatga	taaggtttgt	attottotto	acaaggcctg	ttacctttact	acatytycyg	480
Į	tctgaac	qactcctaaa	ttactcttaa	actatanent	ctgcctttct	catttgtaaa	540
	tataatocto	taccataset	toactetaa	agtgtaaggt	cttccgtgcc	tgtttgtata	600
	taggettete	cgccgcgact	Laccttttgt	accatttgtt	caaatgtatg	gcctggacac	660
	Lagggligic	aaaaatgtct	catgacttca	cccttctttc	ttgtcttggt	gcccgtttta	720
	ttggtctgag	aacggcgtga	tgccaaattc	tgccgctcca	cagtgagcat	ttcgatctac	780
	rggaaattga	ccaacttatt	ttatcacttg	ataactagag	tctgggttca	aacaaaatcc	840
	aataacttca	atcatacatt	gtatttatat	tgaaaaaatt	atgcacaact	Caqtaaatta	900
	cctttttttg	cagtcaaaaa	ttctagatca	gtttctaatt	aatcaaaata	OCCTTTATA	960
	ggtcccagtt	ccattaatat	acctgccgtc	catgcactga	ttacaagaca	Satsoata	
	tatgtttgtt	agtgcttggt	aatataaaac	cttttcttt	2+22=2===	aataccccac	1020
	taattttcta	tttataacat	aactantate		acgagaaagt	tcaccgaaaa	1080
	aattttacaa	ttaaacaca+	annent	gaagtatata (	gtgcagattg	acaagaattt	1140
	aattttgcag -	cryyytatat (	yaacaatttt	cctcaaagtt ( Seite 5	gtagaaaata	tttttcattt	1200

tcttgtcacc	gaaaattatt	tataattgaa	attgaaaccg	aatgagctgc	aagactcgag	1260
tcgaatttca	aaaaaattga	ccaactaaat	atgaaaaaat	ccgaatatat	ccccacccc	1320
ctaccaaaca	catcctaaag	tgaggtatag	actgggactg	ggattgggaa	aaqqqtaaaa	1380
tgctttcact	agcttagcaa	agattccact	ttgttagcta	tctttctttc	tcatttcctt	1440
ttttctttt	cttttttttg	ttatataagc	caaagtaggt	acccaaaagc	atcaatattt	1500
tgtattgctt	ggtgattcct	ctttactcca	gtatttcatt	ttctacaagt	tccacctccc	1560
tccataatta	accattatca	atcttataca	***	3 4		1300
		acceataca	ccccata			1599

<210> 6

<211> 1204

<212> DNA

13> Lycopersicon esculentum

<220>

<400> 6

<221> promoter

<222> (1)..(1204)

<223> B-Gene Promotor

tctgaaaagg agcaccatat ttgccgcact gtggttcata tttccaagta catttagatg 60 aactatatca tcagattgaa aggttattgt ataatcaatc cagtggattc tcgttctggc 120 acctttagaa gtacatgtgc ggaaaagaat gataaggttt gtattgttgt tgacaaagcc 180 tgttgccttt ctcatttgta aatgttctga acgactccta aattactctt aaggtgtaag 240 ttccgtg cctgtttgta aatataatgc tgtgccgtga cttacctttt gtaccatttg 300 aaatgta tggcctggac accagggttg tcaaaaatgt ctcatgcccg ttttattggt 360 ctgagaatgg cgtgatgcca aattctgccg ctccacagtg agcatttcga tctactggaa 420 attgaccaac ttattttatc acttgataac taaacaaaat cctattaact ttaatcatac 480 attgtattta taccgaaaaa gttatgcata actcagtaaa ttaccttttt tagcagtcaa 540 attctagatc agtttctaat ttatcaaaat ggcttttata gggtcccagt tccactaata 600 tacctgccgt ccatgcactg actacaagac aaatacctca ctatgtttgt tagtgcttgg 660 taatataaaa ccttttcttt tatgagaaag ttcaccgaga ataattttct atttgtggca 720 taactagtat atagtgcaga ttgacaagaa tttaattttg cagttgggca catgaacaat 780 tttcctcaaa gttgtagaaa gtacttttca ttttcttgtc accgaaaatt atttataatt 840 gaaattaaaa ccgaatgagc tgcaagattc aagtcgaatt tcaaaagaat tgaccaagaa 900 aaaattcaaa aatatccccc acccctacc aaacacatcc taaagtgagg tatagactgg 960 Seite 6

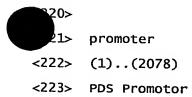
gactgggatt gggaaaaggg taaaatgctt	tcactagctt	agcaaagatt	ccactttgtt	1020
agctatcttt ctttctcatt tcctttttc	tttttcttt	ttttgttata	taagccaaag	1080
taggtaccca aaagcatcaa tattttgtat	tgcttggtga	ttcctctgta	gtccagtatt	1140
tcattttcta caagttccac ctccctccat	aattaaccat	tatcaatctt	atacattctc	1200
tata				1204

<210> 7

<211> 2078

<212> DNA

<213> Lycopersicon esculentum

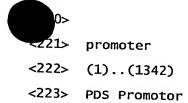


<400> 7	ttacaacaac	****				
cccgccagca	ttacaacagc	ttatatgttg	agcaggtaaa	agcttcaatg	ccctattctt	60
tctacagtta	tcaatgttgc	tcgtctaata	tctggtgttc	ttctcgaaat	gtcaattggc	120
ttgcagcaca	ttgtcctcta	atatccattc	aagcttctta	gatgatgaaa	catttgtcaa	180
atttattaat	ttcatagtgt	tcagtctcaa	ttctttagct	ggttcctcat	agtaaagttg	240
tctaatatga	aatgaaaatg	ttctgtgtgt	tgtactaata	ccttttcatg	gttgtctata	300
gaacgtcgat	gaagagccaa	acagaaacta	ttttgggctg	cgatttctga	taccattgta	360
gaatgct	gggtgggagc	tcatcagaag	ctttacaatg	ggtcacatat	atggagccgg	420
	gctgggaatc					480
tctgcccaca	gcccagttga	ttacgtgaac	tccgtcagac	ttggaaagga	gagaagtacc	540
	ctttttagaa					600
gatcatgcag	aaggcgtcca	gtttagtttt	tgaaggttgt	ttggagttta	tttatctaaa	660
gtaaacttaa	atcagctttt.	tgtttatgag	ttcagtgaac	tatatgttca	aataagactt	720
	atatgtgttt					780
	tagctaccat					840
	agcattcagg					900
	cttcaaatca					960
	gtttttacat					1020
	tctgtaattg			cagcaatgtg		1080

gattttcctt	cagagtagaa	attgaaaaga	atcaactaaa	aaggatagtc	cttcgatttg	1140
atttccggct	taaaaataaa	ctaataagaa	tgagagagcg	aataatagaa	tattttgaaa	1200
ttttaaagat	attcaactat	gttaaattgc	gttataaatt	tcttaaatta	gtagcaccta	1260
atagtttagt	tctcaaaagt	caaaactact	acataatgtg	ctcatttttc	acattaaaat	1320
gcctacatga	tgtaaaagta	aaactcgtag	cattctacgt	gttttactca	actcaaacat	1380
			ttctctctcc			1440
ttttaaaaaa	atatttttt	ttatatcaat	ccaaatgggc	tccaatttat	cataaattag	1500
gtagaaactt	agatattaaa	gaaagaaaag	ggtttatctc	gcaagtgtgg	ctatggtggg	1560
acgtgtcaaa	ttttggattg	tagccaaaca	tgagatttga	tttaaaggga	attggccaaa	1620
tcaccgaaag	caggcatctt	catcataaat	tagtttgttt	atttatacag	aattatacgc	1680
ttttactagt	tatagcattc	ggtatctttt	tctgggtaac	tgccaaacca	ccacaaattt	1740
			ccaaccaaat			1800
aaccactccc	tatatcttct	aggtgctttc	attcgttccg	aggtaagaaa	agatttttgt	1860
ttctttgaat	gctttatgcc	actcgtttaa	cttctgaggt	ttgtggatct	tttaggcgac	1920
ttttttttt	tttgtatgta	aaatttgttt	cataaatgct	tctcaacata	aatcttgaca	1980
aagagaagga	attttaccaa	gtatttaggt	tcagaaatgg	ataattttct	tactgtgaaa	2040
tatccttatg	gcaggtttta	ctgttatttt	tcagtaaa		- 2	2078
					•	_

<210> 8 <211> 1342 <212> DNA

<213> Lycopersicon esculentum



<400>	8						
gtttt	ttttg	ttgttgagca	ctttgtgtgc	attggataaa	ccccaacgt	gtaatagcta	60
ccata	caaga	gaagtaactc	gcactgtcca	tgtcttatgt	ggctcgactc	agaaagcatt	120
caggg	ggatt	gataaccacc	ctccaaacca	actgaaccat	tgtgaataac	cacccttcaa	180
			ggacaaatat				240
acatg	ttcc	tcttacttct	ttagttttct	tgaccatatc	ttgcgttttt	cccttctgta	300
attgad	cactt	ttcttcaaac	catccagcaa	tgtggaagct Seite 8	tgacgatttt	ccttcagagt	360

agaaattgaa	aagaatcaac	taaaaaggat	agtccttcga	tttgatttcc	ggcttaaaaa	420
taaactaata	agaatgagag	agcgaataat	agaatatttt	gaaattttaa	agatattcaa	480
ctatgttaaa	ttgcgttata	aatttcttaa	attagtagca	cctaatagtt	tagttctcaa	540
aagtcaaaac	tactacataa	tgtgctcatt	tttcacatta	aaatgcctac	atgatgtaaa	600
agtaaaactc	gtagcattct	acgtgtttta	ctcaactcaa	acatcctgtt	cattttaata	660
aacgtacgat	gagcttctct	ctccaatttt	cttttctttt	tttttttaa	aaaaatattt	720
ttttttatat	caatccaaat	gggctccaat	ttatcataaa	ttaggtagaa	acttagatat	780
taaagaaaga	aaagggttta	tctcgcaagt	gtggctatgg	tgggacgtgt	caaattttgg	840
attgtagcca	aacatgagat	ttgatttaaa	gggaattggc	caaatcaccg	aaagcaggca	900
tcttcatcat	aaattagttt	gtttatttat	acagaattat	acgcttttac	tagttatagc	960
attcggtatc	tttttctggg	taactgccaa	accaccacaa	atttcaagtt	tccatttaac	1020
tcaactt	caacccaacc	aaatttattt	gcttaattgt	gcagaaccac	tccctatatc	1080
ttctaggtgc	tttcattcgt	tccgaggtaa	gaaaagattt	ttgtttcttt	gaatgcttta	1140
tgccactcgt	ttaacttctg	aggtttgtgg	atcttttagg	cgactttttt	tttttttgta	1200
tgtaaaattt	gtttcataaa	tgcttctcaa	cataaatctt	gacaaagaga	aggaatttta	1260
ccaagtattt	aggttcagaa	atggataatt	ttcttactgt	gaaatatcct	tatggcaggt	1320
tttactgtta	tttttcagta	aa				1342

<210> 9

<211> 1008

<212> DNA

<213> Lycopersicon esculentum



<221> promoter

<222> (1)..(1008)

<223> PDS Promotor

<400> 9		•				
aagcttgacg	attttccttc	agagtagaaa	ttgaaaagaa	tcaactaaaa	aggatagtcc	60
ttcgatttga	tttccggctt	aaaaataaac	taataagaat	gagagagcga	ataatagaat	120
attttgaaat	tttaaagata	ttcaactatg	ttaaattgcg	ttataaattt	cttaaattag	180
·tagcacctaa	tagtttagtt	ctcaaaagtc	aaaactacta	cataatcgtg	ctcatttttc	240
acattaaaat	gcctacatga	tgtaaaagta	aaactcgtag	cattctacgt	gttttactca	300
actcaaacat	cctgttcatt	ttaataaacg	tacgatgagc Seite 9		aattttcttt	360

		atatttttt				420
		agatattaaa				480
		ttttggattg				540
		caggcatctt				600
		tatagcattc				660
		tttaactctt				720
		tatatcttct				780
		gctttatgcc				840
tttaggcgac	tttttttt	tttgtatgta	aaatttgttt	cataaatgct	tctcaacata	900
aatcttgaca	aagagaagga	attttaccaa	gtatttaggt	tcagaaatgg	ataattttct	960
tactgtgaaa	tatccttatg	gcaggtttta	ctgttatttt	tcagtaaa		1008

**10> 10** 

<211> 2078

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> promoter

<222> (1)..(2078)

<223>

<400> gccagta ttacaacagc ttatatgttg agcaggtaaa agcttcaatg ccctattctt 60 acagtta tcaatgttgc tcgtctaata tctggtgttc ttctcgaaat gtcaattggc 120 ttgcagcaca ttgtcctcta atatccattc aagcttctta gatgatgaaa catttgtcaa 180 atttattaat ttcatagtgt tcagtctcaa ttctttagct gtttcctcat agtaaagttg 240 tctaatatga aatgaaaatg ttctgtgtgt tgtactaata ccttttcatg gttgtctata 300 gaacgtcgat gaagagccaa acagaaacta ttttgggctg cgatttctga taccattgta 360 tctgaatgct gggtgggagc tcatcagaag ctttacaatg ggtcacatat atggagccga 420 gtatgaggaa tgctgggaat cagttgtgct tcgcgtgcta ggacttttcc ttcctggtat 480 ttctgcccac agcccagttg attacgtgaa ctccgtcaga cttggaaagg agagaagtac 540 ccaaatgtcg tctttttaga aatacttttg tcacaaaata gcggggttta cagctacaga 600 agatcatgca gaaggcgtcc agtttagttt ttgaaggttg tttggagttt atttatctaa 660 agtaaactta aatcagcttt ttgtttatga gttcagtgaa ctatatgttc aaataagact 720 Seite 10

```
tccctttgta gaatatgtgt ttttttttgt tgttgagcac tttgtgtgca ttggataaac
                                                                      780
ccccaacgtg taatagctac catacaagag aagtaactcg cactgtccat gtcttatgtg
                                                                      840
gctcgactca gaaagcattc agggggattg ataaccaccc tccaaaccaa ctgaaccatt
                                                                      900
gtgaataacc acccttcaaa tcaaccgagt cctcgtgaag gacaaatatg tggttttata
                                                                      960
tacattaaat tttgttttta catgcttcct cttacttctt tagttttctt gaccatatct
                                                                     1020
tctttttccc ttctgtaatt gacattttct tcaaaccatc cagcaatgtg gaagcttgac
                                                                     1080
gattttcctt cagagtagaa attgaaaaga atcaactaaa aaggatagtc cttcgatttg
                                                                     1140
atttccggct taaaaataaa ctaataagaa tgagagagcg aataatagaa tattttgaaa
                                                                     1200
ttttaaagat attcaactat gttaaattgc gttataaatt tcttaaatta gtagcaccta
                                                                     1260
atagtttagt tctcaaaagt caaaactact acataatgtg ctcatttttc acattaaaat
                                                                     1320
gcctacatga tgtaaaagta aaactcgtag cattctacgt gttttactca actcaaacat
                                                                     1380
   gttcatt ttaataaacg tacgatgagc ttctctctcc aattttcttt tcttttttt
                                                                     1440
etttaaaaaa atatttttt ttatatcaat ccaaatgggc tccaatttat cataaattag
                                                                    1500
gtagaaactt agatattaaa gaaagaaaag ggtttatctc gcaagtgtgg ctatggtggg
                                                                    1560
acgtgtcaaa ttttggattg tagccaaaca tgagatttga tttaaaggga attggccaaa
                                                                    1620
tcaccgaaag caggcatctt catcataaat tagtttgttt atttatacag aattatacgc
                                                                    1680
ttttactagt tatagcattc ggtatctttt tctgggtaac tgccaaacca ccacaaattt
                                                                    1740
caagtttcca tttaactctt caacttcaac ccaaccaaat ttatttgctt aattgtgcag
                                                                    1800
aaccactccc tatatcttct aggtgctttc attcgttccg aggtaagaaa agattttgt
                                                                    1860
ttctttgaat gctttatgcc actcgtttaa cttctgaggt ttgtggatct tttaggcgac
                                                                    1920
ttttttttt tttgtatgta aaatttgttt cataaatgct tctcaacata aatcttgaca
                                                                    1980
aagagaagga attttaccaa gtatttaggt tcagaaatgg ataattttct tactgtgaaa
                                                                    2040
  ccttatg gcaggtttta ctgttatttt tcagtaaa
                                                                    2078
```

<210> 11

<211> 1528

<212> DNA

<213> Cucumis sativus

<220>

<221> promoter

<222> (1)..(1528)

<223> CHRC Promotor

<400> 11			043equ.L	ХL		
	g ggttacttta	ttcattttca	tccattctct	ttattgttaa	attttgtaca	60
tttattcaa	t aatattatat	gtttattaca	aattctcact	ttcttattca	tacctattca	120
ctcaagcct	t taccatcttc	cttttctatt	tcaatactat	ttctacttca	tttttcacgt	180
ttttaacat	c tttctttatt	tcttgtccac	ttcgtttagg	gatgcctaat	gtcccaaatt	240
tcatctctc	g tagtaacaca	aaaccaatgt	aatgctactt	ctctctacat	ttttaataca	300
aataaagtg	a aacaaaatat	ctataaataa	acaaatatat	atattttgtt	agacgctgtc	360
tcaacccat	c aattaaaaaa	ttttgttata	tttctacttt	acctactaaa	tttgtttctc	420
atatttacc	t tttaaccccc	acaaaaaaaa	attataaaaa	agaaagaaaa	aagctaaacc	480
ctatttaaa	t agctaactat	aagatcttaa	aattatcctc	atcagtgtat	agtttaattg	540
gttattaac	t tataacatta	tatatctatg	acatatactc	tctcctagct	atttctcaca	600
ttttttaac	t taagaaaata	gtcataacat	agtctaaaat	tcaaacatcc	acatgctcta	660
tgatta	a caaaaagtta	gaaatattta	tttaaataaa	aaagactaat	aaatatataa	720
aatgaatgt	t catacgcaga	cccatttaga	gatgagtatg	ctttcacatg	ctgagattat	780
tttcaaaac	t aaggttgtag	caatattaaa	tcaataaaat	tattataaat	aacaaaatta	840
acctgctcg	t gtttgctgta	tatgggaggc	tacaaaataa	attaaactaa	agatgattat	900
gttttagac	a ttttttctat	ctgtattagt	ttatacatat	taattcagga	gctgcacaac	960
ccaattcta	t tttcgttcct	tggtggctgg	gtttctcaca	aggttcaata	gtcaatatta	1020
ggttttatt	g gacttttaat	agtatcaaac	aaatctatgt	gtgaacttaa	aaattgtatt	1080
aaatattta	g ggtaacctgt	tgccgttttt	agaataatgt	ttcttcttaa	tacacgaaag	1140
cgtattgtg	t attcattcat	ttggcgcctc	acatgcttcg	gttggctcgc	tttagtctct	1200
gccttcttt	g tatattgtac	tcccctctt	cctatgccac	gtgttctgag	cttaacaagc	1260
cacgttgcg	t gccattgcca	aacaagtcat	tttaacttca	caaggtccga	tttgacctcc	1320
aacaacg	a caagtttccg	aacagtcgcg	aagatcaagg	gtataatcgt	ctttttgaat	1380
atttct	c tttatttaat	agtccctctc	gtgtgatagt	ttttaaaaga	tttttaaaac	1440
gtagctgct	g tttaagtaaa	tcccagtcct	tcagtttgtg	cttttgtgtg	ttttgtttct	1500
ctgatttac	g gaatttggaa	ataattct				1528

<210> 12

<211> 1538

<212> DNA

<213> Cucumis sativus

<220>

<221> promoter

<222> (1)..(1538) <223> CHRC Promotor

	<400> 12						
						attttgtaca	60
						tacctattca	120
	ctcaagcctt	taccatcttc	cttttctatt	tcaatactat	ttctacttca	tttttcacgt	180
		tttctttatt					240
		tagtaacaca					300
		aacaaaatat					360
		aattaaaaaa					420
		tttaaccccc					480
		agctaactat					540
		tataacatta					600
	ttttttaact	taagaaaata	gtcataacat	agtctaaaat	tcaaacatcc	acatgctcta	660
		caaaaagtta					720
		catacgcaga					780
		aaggttgtag					840
		gtttgctgta					900
	gttttagaca	ttttttctat	ctgtattagt	ttatacatat	taattcagga	gctgcacaac	960
	ccaattctat	tttcgttcct	tggtggctgg	gtttctcaca	aggttcaata	gtcaatatta	1020
1	ggttttattg	gacttttaat	agtatcaaac	aaatctatgt	gtgaacttaa	aaattgtatt	1080
į	aaatatttag	ggtaacctgt	tgccgttttt	agaataatgt	ttcttcttaa	tacacgaaag	1140
		attcattcat					1200
		tattttgtac					1260
(	cacgttgcgt	gccattgcca	aacaagtcat	tttaacttca	caaggtccga	tttgacctcc	1320
ä	aaaacaacga	caagtttccg	aacagtcgcg	aagatcaagg	gtataatcgt	ctttttgaat	1380
		tttatttaat					1440
		tttaagtaaa					1500
		gaatttggaa				•	1538

<210> 13

<211> 1525

<212> DNA

<213> Cucumis sativus

<221> promoter <222> (1)..(1525) <223> CHRC Promotor <400> 13 tacaaattag ggttacttta ttcattttca tccattctct ttattgttaa attttgtaca 60 tttattcaat aatattatat gtttattaca aattctcact ttcttattca tacctattca 120 ctcaagcctt taccatcttc cttttctatt tcaatactat ttctacttca tttttcacgt 180 ttttaacatc tttctttatt tcttgtccac ttcgtttagg gatgcctaat gtcccaaatt 240 tcatctctcg tagtaacaca aaaccaatgt aatgctacct ctctctacat ttttaataca 300 aaagtga aacaaaatat ctataaataa acaaatatat atattttgtt agacgctgtc 360 tcaacccatc aattaaaaaa ttttgttata tttctacttt acctactaaa tttgtttctc 420 atatttacct tttaaccccc acaaaaaaa attataaaaa agaaagaaaa aagctaaacc 480 ctatttaaat agctagctat aagatcttaa aattatcctc atcagtgtat agtttaattg 540 gttattaact tataacatta tatatctatg acatatactc tctcctagct atttctcaca 600 ttttttaact taagaaaata gtcataacat agtctaaaat tcaaacatcc acatgctcta 660 atttgattaa caaaaagata gaaatattta tttaaataaa aaagactaat aaatatataa 720 aatgaatgtt catacgcaga cccatttaga gatgagtatg ctttcacatg ttgagattat 780 tttcaaaact aaggttgtag caatattaaa tcaataaaat tattataaat aacaaaatta 840 acctgctcgt gtttgctgaa tatgggaggc tacaaaataa attaaactaa agatgattat 900 gttttagaca ttttttctat ctgtattagt ttatacatat taattcagga gctgcacaac 960 attctat tttcgttcct tggtggctgg gtttctcaca aggttcaata gtcaatatta 1020 tttattg gacttttaat agtatcaaat aaatctatgt gtgaacttaa aaattgtatt 1080 aaatatttag ggtaacctgt tgccgttttt agaataatgt ttcttcttaa tacacgaaag 1140 cgtattgtgt attcattcat ttggcgcctc acatgcttcg gttggctcgc tttagtctct 1200

gccttctttg tattttgtac tcccctctt cctatgccac gtgttctgag cttaacaagc

cacgttgcgt gccattgcca aacaagtcat tttaacttca caaggtccga tttgacctcc

aaaacaacga caagtttccg aacagtcgcg aagatcaagg gtataatcgt ctttttgaat

tctatttctc tttatttaat agtccctctc gtgtgatagt tttttaaaga tttttaaaac

gtagctgctg tttaagtaaa tcccagtcct tcagtttgtg cttttgtgtg ttttgttct

<210> 14

ctgatttacg gaatttggaa ataat

<220>

1260

1320

1380

1440

1500

<211> 1519
<212> DNA
<213> Cucumis sativus

<220>
<221> promoter
<222> (1)..(1519)
<223> CHRC Promotor

<400> 14 ttactttatt cattttcatc cattctcttt attgttaaat tttgtacatt tattcaataa 60 tattatatgt ttattacaaa ttctcacttt cttattcata cctattcact caagccttta 120 tcttcct tttctatttc aatactattt ctacttcatt tttcacgttt ttaacatctt 180 240 ctttatttc ttgtccactt cgtttaggga tgcctaatgt cccaaatttc atctctcgta 300 gtaacacaaa accaatgtaa tgctacctct ctctacattt ttaatacaaa taaagtgaaa 360 caaaatatct ataaataaac aaatatatat attttgttag acgctgtctc aacccatcaa 420 ttaaaaaatt ttgttatatt tctactttac ctactaaatt tgtttctcat atttaccttt 480 taacccccac aaaaaaaaat tataaaaaag aaagaaaaaa gctaaaccct atttaaatag ctagctataa gatcttaaaa ttatcctcat cagtgtatag tttaattggt tattaactta 540 600 taacattata tatctatgac atatactctc tcctagctat ttctcacatt ttttaactta 660 agaaaatagt cataacatag tctaaaattc aaacatccac atgctctaat ttgattaaca 720 aaaagataga aatatttatt taaataaaaa agactaataa atatataaaa tgaatgttca tacgcagacc catttagaga tgagtatgct ttcacatgtt gagattattt tcaaaactaa 780 840 ttgtagca atattaaatc aataaaatta ttataaataa caaaattaac ctgctcgtgt ctgaata tgggaggcta caaaataaat taaactaaag atgattatgt tttagacatt 900 ttttctatct gtattagttt atacatatta attcaggagc tgcacaaccc aattctattt 960 tcgttccttg gtggctgtgt ttctcacaag gttcaatagt caatattagg ttttattgga 1020 cttttaataq tatcaaataa atctatgtgt qaacttaaaa attgtattaa atatttaggg 1080 taacctqttq ccgtttttag aataatqttt cttcttaata cacgaaagcg tattgtgtat 1140 tcattcattt ggcgcctcac atgcttcggt tggctcgctt tagtctctgc cttctttgta 1200 ttttgtactc cccctcttcc tatgccacgt gttctgagct taacaagcca cgttgcgtgc 1260 cattgccaaa caagtcattt taacttcaca aggtccgatt tgacctccaa aacaacgaca 1320 1380 agtttccgaa cagtcgcgaa gatcaagggt gtaatcgtct ttttgaattc tatttctctt 1440 tatttaatag tccctctcgt gtgatagttt tttaaagatt tttaaaacgt agctgctgtt 1500 taagtaaatc ccagtccttc agtttgtgct tttgtgtgtt ttgtttctct gatttacgga

Seite 15

att	ttgga	aaat	aata	agct	tt											1519
<21	L0>	15														
<21	L1>	1773	Ĺ													
<21	L2>	DNA														
<21	L3>	Haen	natoc	occu	ıs pl	uvia	lis									
<22																
<22		CDS														
<22		(166	5)(	1155	<b>)</b> .											
<22	:3>		•													
	0>	15														
			tgca	cgca	ag t	cagc	gcgc	g ca	agto	aaca	cct	gccg	gtc	caca	gcctca	60
															ttgaac	120
												ga a	tg c	ag c	ta gca	177
			•									1			eu Ala	
gcg Ala	aca Thr	gta Val	atg Met	ttg Leu	gag Glu	cag Gln	ctt Leu	acc Thr	gga Glv	ago	gct	gag	gca	ctc	aag Lys	225
					10					ΤЭ					20	
Glu	Lys	gag Glu	aag Lys		gtt Val	gca Ala	ggc Gly	agc Ser	261	gac Asp	gtg Val	ttg Leu	cgt Arg	aca Thr	tgg Trp	273
									30					35		
Ala	Thr	Gln	Tyr 40	Ser	Leu	Pro	Ser	Giu	gag Glu	tca Ser	gac Asp	gcg Ala	gcc Ala	cgc Arg	ccg Pro	321
3								43					50			
	Leu	Lys 55	Asn	Āla	Tyr	Lys	Pro 60	cca Pro	Pro	Ser	gac Asp	Inr	aag Lys	ggc Gly	atc Ile	369
aca	atg	gcg	cta	cgt	gtc	atc	aac	tcc	taa	acc	uca	65 ata	++c	ctc	62.6	417
Thr	Met 70	Ala	Leu	Arg	Va 1	11e 75	ĞÎy	Ser	Trp	Āla	Ala 80	Val	Phe	Leu	His	417
gcc	att	ttt	caa	atc	aag	ctt	ccg	acc	tcc	tta	~~~	cao	cta	cac	taa	465
85 85	TIE	Pne	GIN	Ile	Lys 90	Leu	Pro	Thr	Ser	Leŭ 95	Ăsp	Gl'n	Leu	His	Trp 100	703
ctg	CCC	gtg	tca	gat	gçc	aca	gct	cag	ctg	gtţ	agc.	ggc	acg	agc		513
		Vai	261	105	Ald	Inr	AIA	Gln	Leu 110	Val	Ser	Gly	Thr	Ser 115	Ser	
ctg Leu	ctc Leu	gac Asn	atc	gtc Val	gta	gta	ttc	ttt Phe	gtc	ctg	gag	ttc	ctg	tac	aca	561
		, p	120	·ω,	vu.	vai	rne	125	vai	Leu	GIU	Pne	<b>Leu</b> 130	Tyr	Thr	
ggc Gly	ctt Leu	ttt Phe	atc Ile	acc Thr	acg Thr	cat His	gat Asp	gct Ala	atg Met	cat	ggc	acc	atc	gcc	atg	609
		135				_	140		_	1113 Fe 16		145	T 16	AId	MEC	
									~ - I	- 11	•					

										-						
aga Arg	aac Asn 150	agg Arg	cag Gln	ctt Leu	aat Asn	gac Asp 155	ttc Phe	ttg Leu	ggc Gly	aga Arg	gta Val 160	tgc Cys	atc Ile	tcc Ser	ttg Leu	657
165	,,,,			ASP	170	A311	MEL	Leu	ніѕ	175	aag Lys	HIS	Trp	Glu	His 180	705
				-05					190		gac Asp			195		753
		,	200	vui	710	пр	FILE	205	ser	Pne	atg Met	Ser	Ser 210	Tyr	Met	801
		215			,,,,	<u>g</u>	220	AIQ	ΠÞ	πρ	acg Thr	225	vaı	Met	Gin	849
	230	- · <b>J</b>			1100	235	MOII	Leu	Leu	Vai	ttc Phe 240	мет	Ala	Ala	Ala	897
Pro 245	atc Ile	ctg Leu	tcc Ser	gcc Ala	ttc Phe 250	cgc Arg	ttg Leu	ttc Phe	tac Tyr	ttt Phe 255	ggc Gly	acg Thr	tac Tyr	atg Met	ccc Pro 260	945
	-,-			265	ary	AIA	міа	261.	270	ser	tca Ser	Pro	Ala	Va 1 275	Met	993
	•		280		/ y	••••	361	285	AId	ser	gac Asp	Leu	va 1 290	Ser	Phe	1041
ctg Leu	acc Thr	tgc Cys 295	tac Tyr	cac His	ttc Phe	gac Asp	ctg Leu 300	cac His	tgg Trp	gag Glu	cac His	cac His 305	cgc Arg	tgg Trp	CCC Pro	1089
ttc Phe	gcc Ala 310	ccc Pro	tgg Trp	tgg Trp	gag Glu	ctg Leu 315	ccc Pro	aac Asn	tgc Cys	cgc Arg	cgc Arg 320	ctg Leu	tct Ser	ggc Gly	cga Arg	1137
	ctg Leu	gtt Val	cct Pro	gcc Ala	tag	ctgg	acac	ac t	gcag	tggg	c cc	tgct	gcca			1185
gctg	ggca	tg c	aggt	tgtg	g ca	ggac	tggg	tga	ggtg	aaa	agct	gcag	gc g	ctgc	tgccg	1245
gaca	cgct	gc a	tggg	ctac	c ct	gtgt	agct	gcc	gcca	cta	gggg	aggg	gg t	ttgt	agctg	1305
tcga	gctt	gc c	ccat	ggat	g aa	gctg	tgta	gtg	gtgc	agg	gagt	acac	cc a	cagg	ccaac	1365
accc	ttgc	ag g	agat	gtct	t gc	gtcg	ggag	gag	tgtt	ggg	cagt	gtag	at g	ctat	gattg	1425
tatc	ttaa	tg c	tgaa	gcct	t ta	gggg	agcg	aca	ctta	gtg	ctgg	gcag	gc a	acgc	cctgc	1485
aagg	tgca	gg c	acaa	gcta	g gc	tgga	cgag	gac	tcgg	tgg	cagg	cagg	tg a	agag	gtgcg	1545
															cagtg	1605
agag	ctgc	gt g	atta	actg	g gc	tatg	gatt	gtt	tgag	cag	tctc	actt	at t	cttt	gatat	1665
agat	actg	gt ca	aggc	aggt	c ag	gaga	gtga	gta	tgaa	caa	gttg	agag	gt g	gtgc	gctgc	1725
ccct	gcgc <sup>.</sup>	tt a	tgaa	gctg <sup>.</sup>	t aa	caata	aaag		ttca: Seit			aa				1771

<210> 16

<211> 329

<212> PRT

<213> Haematococcus pluvialis

<400> 16

Met Gln Leu Ala Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala 10 15

Glu Ala Leu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val 20 25 30

eu Arg Thr Trp Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp 35 40 45

Ala Ala Arg Pro Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp 50 60

Thr Lys Gly Ile Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala 65 70 75 80

Val Phe Leu His Ala Ile Phe Gln'Ile Lys Leu Pro Thr Ser Leu Asp 85 90 95

Gln Leu His Trp Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser 100 105 110

Gly Thr Ser Ser Leu Leu Asp Ile Val Val Phe Phe Val Leu Glu 115 120 125

Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly 130

Thr Ile Ala Met Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val 150 155 160

Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys 165 170 175

His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp 180 185 190

Phe His Arg Gly Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met 200 205

Ser Ser Tyr Met Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr 210 220 Seite 18

# O4sequ.txt

Val Val Met Gln Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe 225 230 235 240	
Met Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly 245 250 255	
Thr Tyr Met Pro His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser 260 265 270	
Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp 275 280 285	
Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His 290 295 300	
s Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg 310 315 320	
Leu Ser Gly Arg Gly Leu Val Pro Ala 325	
<210> 17	
<211> 1662	
<212> DNA .	
<213> Haematococcus pluvialis	
<220>	
<222> (168)(1130) 3>	
<400> 17 cggggcaact caagaaattc aacagctgca agcgcgcccc agcctcacag cgccaagtga	60
gctatcgacg tggttgtgag cgctcgacgt ggtccactga cgggcctgtg agcctctgcg	120
ctccgtcctc tgccaaatct cgcgtcgggg cctgcctaag tcgaaga atg cac gtc Met His Val	176
gca tcg gca cta atg gtc gag cag aaa ggc agt gag gca gct gct tcc Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala Ala Ala Ser 5 10 15	224
agc cca gac gtc ttg aga gcg tgg gcg aca cag tat cac atg cca tcc Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His Met Pro Ser 20 25 30 35	272
gag tcg tca gac gca gct cgt cct gcg cta aag cac gcc tac aaa cct Seite 19	320

	G]	11 SA	r Se	n A.	m A T.	7	_		_	04s	equ.	txt					
										a Le 45	u Ly:	s Hi			50	s Pro	
				55				,	60	ME	L Ale	a Lei	ומו ג	65	2 I I	t ggc e Gly	368
			70					75	, 410		= PNE	e Gir	1 116 80	e Arg	) Lei	a ccg u Pro	416
	aca Thi	s tco Ser 85	c at	g ga t As <sub>l</sub>	c cag p Glr	cti Lei	cac His 90	tgg Trp	tto Lei	CC1	gto Val	tco Ser 95	gaa Glu	gco I Ala	aca Thi	a gcc Ala	464
	100	)		-	-	105	5			Let	110	)	: Ala	LAIa	va i	ttc Phe 115	512
_	att Ile	gta Val	a cti Lei	gaç Gli	tto Phe 120	ctg Leu	tac Tyr	act Thr	ggt	Cta Leu 125	PHE	atc lle	acc Thr	aca Thr	cat His	gac Asp	560
				135	•			LCu	140	1113	Arg	GIN	Leu	Asn 145	Asp	ctc	608
			150	)	tgc Cys			155	1 9 1	АТА	πρ	Pne	160	Туг	Ser	Met	656
		165			cac His	•	170		,,,,	7311	піз	175	GIY	GIU	vaı	Gly	704
	180				ttc Phe	185	, -	,	,,,,,,		190	Leu	Vai	Pro	ırp	Phe 195	752
					tcc Ser 200		.,.	1-10-6	361	205	пр	GIN	Pne	Ala	Arg 210	Leu	800
Ű			•	215	gtg Vai			G ( ) ,	220	Leu	Gry	Ala	Pro	Met 225	Ala	Asn	848
			230		atg Met			235			LCu	261	240	Pne	Arg	Leu	896
		245		•	act Thr	•	250			Lys	PIO	255	Pro <sub>.</sub>	GIY	Pro	Ala	944
	260				gtg Val	265		р		~ı 9	270	Lys	inr	ser	GIU	A1a 275	992
					agt Ser 280				<b>-</b>	285	піз	rne	ASP	Leu	нт s 290	Trp	1040
				295	tgg Trp				300	пр	пр	GIN	Leu	305	cac His	tgc Cys	1088
	cgc	cgc	ctg	tcc	ggg	cgt	ggc	ctg	gtg	cct Seit	gcc e 20	ttg	gca	tga			1130

# O4sequ.txt Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala 310 315 320

cctggtccct ccgct	ggtga cccagcgtct	gcacaagagt	gtcatgctac	agggtgctgc	1190
ggccagtggc agcgc	agtgc actctcagco	: tgtatggggc	taccgctgtg	ccactgagca	1250
ctgggcatgc cactga					1310
ggcgtgctac tgacaa	atggg cgtgctactg	gggtctggca	gtggctagga	tggagtttga	1370
tgcattcagt agcggf	tggcc aacgtcatgt	ggatggtgga	agtgctgagg	ggtttaggca	1430
gccggcattt gagagg					1490
acagccaggg aaatco	ccttc gagagtgatt	atgggacact	tgtattggtt	tcgtgctatt	1550
gttttattca gcagca	agtac ttagtgaggg	tgagagcagg	gtggtgagag	tggagtgagt	1610
gagtatgaac ctggto	cagcg aggtgaacag	cctgtaatga	atgactctgt	ct	1662

210> 18

11> 320

<212> PRT

<213> Haematococcus pluvialis

<400> 18

Met His Val Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala 1 10 15

Ala Ala Ser Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His 20 25 30

Met Pro Ser Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala 35 40 45

Lys Pro Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr 50 . 60

Ile Ile Gly Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile 70 75 80

Arg Leu Pro Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu 90 95

Ala Thr Ala Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala 100 105 110

Ala Val Phe Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr 115 125

Thr His Asp Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu 130 140 Seite 21

Asn Asp Leu Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp 145 150 150 160

Tyr Ser Met Leu His Arg Lys His Trp Glu His His Asn His Thr Gly 165 170 175

Glu Val Gly Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val 180 185 190

Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe 195 200 205

Ala Arg Leu Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro 210 215 220

Met Ala Asn Leu Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala 230 235 240

Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro 245 250 255

Gly Pro Ala Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr 260 265 270

Ser Glu Ala Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp 275 280 285

Leu His Trp Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu 290 295 300

Pro His Cys Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala 305 310 315 320

10> 19

211> 729

<212> DNA

<213> Agrobacterium aurantiacum

<220>

<221> CDS

<222> (1)..(729)

<223>

Met 1	Ser	Ala	His	Ala 5	Leu	Pro	Lys	Ala	Asp 10	Leu	Thr	Ala	Thr	Ser 15	Leu	
atc Ile	gtc Val	tcg Ser	ggc Gly 20	ggc Gly	atc Ile	atc Ile	gcc Ala	gct Ala 25	tgg Trp	ctg Leu	gcc Ala	ctg Leu	cat His 30	gtg Val	cat His	96
gcg Ala	ctg Leu	tgg Trp 35	ttt Phe	ctg Leu	gac Asp	gca Ala	gcg Ala 40	gcg Ala	cat His	ccc Pro	atc Ile	ctg Leu 45	gcg Ala	atc Ile	gca Ala	144
ASII	50	Leu	ч	Leu	acc Thr	55 55	Leu	Ser	Vai	Gly	Leu 60	Phe	Ile	Ile	Alā	192
cat His 65	gac Asp	gcg Ala	atg Met	cac His	ggg Gly 70	tcg Ser	gtg Val	gtg Val	ccg Pro	999 Gly 75	cgt Arg	ccg Pro	cgc Arg	gcc Ala	aat Asn 80	240
gcg Ala	gcg Ala	atg Met	ggc Gly	cag Gln 85	ctt Leu	gtc Val	ctg Leu	tgg Trp	ctg Leu 90	tat Tyr	gcc Ala	gga Gly	ttt Phe	tcg Ser 95	tgg Trp	288
g g	aag Lys	atg Met	atc Ile 100	gtc Val	aag Lys	cac His	atg Met	gcc Ala 105	cat His	cac His	cgc Arg	cat His	gcc Ala 110	gga Gly	acc Thr	336
∆2þ	ASP	115	F10	ASP	ttc Phe	ASP	120	GIY	GIY	Pro	Vai	Arg 125	Trp	Tyr	Ala	384
cgc Arg	ttc Phe 130	atc Ile	ggc Gly	acc Thr	tat Tyr	ttc Phe 135	ggc Gly	tgg Trp	cgc Arg	gag Glu	999 Gly 140	ctg Leu	ctg Leu	ctg Leu	CCC Pro	432
145	-10	Vai	****	vai	tat Tyr 150	Ala	Leu	тіе	Leu	155	ASP	Arg	Trp	Met	Tyr 160	480
<b>vu</b> .	Vai	FIIC	ιιρ	165	ctg Leu	Pro	ser	TIE	170	АІа	Ser	Ile	GIn	Leu 175	Phe	528
gtg Val	ttc Phe	ggc Gly	acc Thr 180	tgg Trp	ctg Leu	ccg Pro	cac His	cgc Arg 185	ccc Pro	ggc Gly	cac His	gac Asp	gcg Ala 190	ttc Phe	ccg Pro	576
ASP	cgc Arg	cac His 195	aat Asn	gcg Ala	cgg Arg	tcg Ser	tcg Ser 200	cgg Arg	atc Ile	agc Ser	gac Asp	ccc Pro 205	gtg Val	tcg Ser	ctg Leu	624
ctg Leu	acc Thr 210	tgc Cys	ttt Phe	cac His	ttt Phe	ggc Gly 215	ggt Gly	tat Tyr	cat His	cac His	gaa Glu 220	cac His	cac His	ctg Leu	сас His	672
Pro 225	acg Thr	gtg Val	ccg Pro	tgg Trp	tgg Trp 230	cgc Arg	ctg Leu	ccc Pro	agc Ser	acc Thr 235	cgc Arg	acc Thr	aag Lys	ggg Gly	gac Asp 240	720
acc Thr	gca Ala	tga														729

<210> 20

<211> 242

<212> PRT

<213> Agrobacterium aurantiacum

<400> 20

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu 10 15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His 20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala 35 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala 50 60

s Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp 85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr 100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala 115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Pro 130 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr 150 155 160

val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe 165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro 180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu 195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His 210 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp 225 230 235 240 Thr Ala

00> 21 tgcaggccg ggcccggtgg ccaatggtcg caaccggcag gactggaaca ggacggcggg 60 ccggtctagg ctgtcgccct acgcagcagg agtttcgg atg tcc gga cgg aag cct Met Ser Gly Arg Lys Pro 1 116 ggc aca act ggc gac acg atc gtc aat ctc ggt ctg acc gcc gcg atc Gly Thr Thr Gly Asp Thr Ile Val Asn Leu Gly Leu Thr Ala Ala Ile 10 15 20 164 ctg ctg tgc tgg ctg gtc ctg cac gcc ttt acg cta tgg ttg cta gat Leu Leu Cys Trp Leu Val Leu His Ala Phe Thr Leu Trp Leu Leu Asp 25 30 35 212 gcg gcc gcg cat ccg ctg ctt gcc gtg ctg tgc ctg gct ggg ctg acc Ala Ala Ala His Pro Leu Leu Ala Val Leu Cys Leu Ala Gly Leu Thr 260 tgg ctg tcg gtc ggg ctg ttc atc atc gcg cat gac gca atg cac ggg Trp Leu Ser Val Gly Leu Phe Ile Ile Ala His Asp Ala Met His Gly 308 gtg gtg ccg ggg cgg ccg cgc gcc aat gcg gcg atc ggg caa ctg ser Val Val Pro Gly Arg Pro Arg Ala Asn Ala Ala Ile Gly Gln Leu 75 80 80 356 gcg ctg tgg ctc tat gcg ggg ttc tcg tgg ccc aag ctg atc gcc aag Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp Pro Lys Leu Ile Ala Lys 90 95 100 404 cac atg acg cat cac cgg cac gcc ggc acc gac aac gat ccc gat ttc His Met Thr His His Arg His Ala Gly Thr Asp Asn Asp Pro Asp Phe 105 115452 ggt cac gga ggg ccc gtg cgc tgg tac ggc agc ttc gtc tcc acc tat Gly His Gly Gly Pro Val Arg Trp Tyr Gly Ser Phe Val Ser Thr Tyr 120 130 500 ttc ggc tgg cga gag gga ctg ctg cta ccg gtg atc gtc acc acc tat Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro Val Ile Val Thr Thr Tyr 135 140 145 548 gcg ctg atc ctg ggc gat cgc tgg atg tat gtc atc ttc tgg ccg gtc 596 Seite 25

```
04sequ.txt
Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr Val Ile Phe Trp Pro Val 155 160 165
ccg gcc gtt ctg gcg tcg atc cag att ttc gtc ttc gga act tgg ctg
Pro Ala Val Leu Ala Ser Ile Gln Ile Phe Val Phe Gly Thr Trp Leu
170 180
                                                                                        644
ccc cac cgc ccg gga cat gac gat ttt ccc gac cgg cac aac gcg agg
Pro His Arg Pro Gly His Asp Asp Phe Pro Asp Arg His Asn Ala Arg
185 190 195
                                                                                        692
    acc ggc atc ggc gac ccg ttg tca cta ctg acc tgc ttc cat ttc
Thr Gly Ile Gly Asp Pro Leu Ser Leu Leu Thr Cys Phe His Phe
200 210
                                                                                        740
ggc ggc tat cac cac gaa cat cac ctg cat ccg cat gtg ccg tgg tgg
Gly Gly Tyr His His Glu His His Leu His Pro His Val Pro Trp Trp
215 220 225 230
                                                                                        788
cgc ctg cct cgt aca cgc aag acc gga ggc cgc gca tga cgcaattcct
Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly Arg Ala
235 240
                                                                                        837
  ttgtcgtg gcgacagtcc tcgtgatgga gctgaccgcc tattccgtcc accgctggat
                                                                                        897
tatgcacggc cccctaggct ggggctggca caagtcccat cacgaagagc acgaccacgc
                                                                                        957
gttggagaag aacgacctct acggcgtcgt cttcgcggtg ctggcgacga tcctcttcac
                                                                                       1017
cgtgggcgcc tattggtggc cggtgctgtg gtggatcgcc ctgggcatga cggtctatgg
                                                                                       1077
gttgatctat ttcatcctgc acgacgggct tgtgcatcaa cgctggccgt ttcggtatat
                                                                                       1137
tccgcggcgg ggctatttcc gcaggctcta ccaagctcat cgcctgcacc acgcggtcga
                                                                                       1197
ggggcgggac cactgcgtca gcttcggctt catctatgcc ccacccgtgg acaagctgaa
                                                                                       1257
gcaggatctg aagcggtcgg gtgtcctgcg cccccaggac gagcgtccgt cgtgatctct
                                                                                       1317
gatcccggcg tggccgcatg aaatccgacg tgctgctggc aggggccggc cttgccaacg
                                                                                       1377
gactgatcgc gctggcgatc cgcaaggcgc ggcccgacct tcgcgtgctg ctgctggacc
                                                                                       1437
gtgcggcggg cgcctcggac gggcatactt ggtcctgcca cgacaccgat ttggcgccgc
                                                                                       1497
  tggctgga ccgcctgaag ccgatcaggc gtggcgactg gcccgatcag gaggtgcggt
                                                                                       1557
  ccagacca ttcgcgaagg ctccgggccg gatatggctc gatcgacggg cgggggctga
                                                                                       1617
tgcgtgcggt gacc
                                                                                       1631
<210>
         22
<211>
         242
<212>
         PRT
<213> Alcaligenes sp.
<400> 22
Met Ser Gly Arg Lys Pro Gly Thr Thr Gly Asp Thr Ile Val Asn Leu
1 10 15
```

Gly Leu Thr Ala Ala Ile Leu Leu Cys Trp Leu Val Leu His Ala Phe 20 25 30 Thr Leu Trp Leu Leu Asp Ala Ala Ala His Pro Leu Leu Ala Val Leu 35 40 45 Cys Leu Ala Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala 50 60 His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn 65 70 75 80 Ala Ala Ile Gly Gln Leu Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp 85 90 95 Pro Lys Leu Ile Ala Lys His Met Thr His His Arg His Ala Gly Thr 100 105 110 p Asn Asp Pro Asp Phe Gly His Gly Gly Pro Val Arg Trp Tyr Gly 115 120 125 Ser Phe Val Ser Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro 130 140 Val Ile Val Thr Thr Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr 145 150 . 155 160 Val Ile Phe Trp Pro Val Pro Ala Val Leu Ala Ser Ile Gln Ile Phe 165 170 175 Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Asp Phe Pro 180 185 190 Asp Arg His Asn Ala Arg Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu 195 200 205 Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His 210. 220 Pro His Val Pro Trp Trp Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly 235 240

Arg Ala

<210> 23

<211> 729

<212> DNA

<213> Paracoccus marcusii

<220>

<221	> C	DS															
<222	> (	1)	(729	)													
<223	>																
<400 atg Met 1	agc	gca	cat His	gcc Ala 5	ctg Leu	ccc Pro	aag Lys	gca Ala	gat Asp 10	ctg Leu	acc Thr	gcc Ala	aca Thr	agc Ser 15	ctg Leu		48
atc Ile	gtc Val	tcg Ser	ggc Gly 20	ggc Gly	atc Ile	atc Ile	gcc Ala	gca Ala 25	tgg Trp	ctg Leu	gcc Ala	ctg Leu	cat His 30	gtg Val	cat His		96
gcg la	ctg Leu	tgg Trp 35	ttt Phe	ctg Leu	gac Asp	gcg Ala	gcg Ala 40	gcc Ala	cat His	ccc Pro	atc Ile	ctg Leu 45	gcg Ala	gtc Val	gcg Ala	;	144
aat Asn	ttc Phe 50	ctg Leu	gjy ggg	ctg Leu	acc Thr	tgg Trp 55	ctg Leu	tcg Ser	gtc Val	gga Gly	ttg Leu 60	ttc Phe	atc Ile	atc Ile	gcg Ala		192
cat His 65	gac Asp	gcg Ala	atg Met	cac His	ggg Gly 70	tcg Ser	gtc Val	gtg Val	ccg Pro	ggg Gly 75	cgt Arg	ccg Pro	cgc Arg	gcc Ala	aat Asn 80		240
											gcc Ala						288
cgc Arg	aag Lys	atg Met	atc Ile 100	gtc Val	aag Lys	cac His	atg Met	gcc Ala 105	cat His	cac His	cgc Arg	cat His	gcc Ala 110	gga Gly	acc Thr		336
gac Asp	gac Asp	gac Asp 115	cca Pro	gat Asp	ttc Phe	gac Asp	cat His 120	ggc Gly	ggc Gly	ccg Pro	gtc val	cgc Arg 125	tgg Trp	tac Tyr	gcc Ala		384
	ttc Phe 130	atc Ile	ggc Gly	acc Thr	tat Tyr	ttc Phe 135	ggc Gly	tgg Trp	cgc Arg	gag Glu	ggg Gly 140	ctg Leu	ctg Leu	ctg Leu	CCC Pro		432
gtc Val 145	atc Ile	gtg Val	acg Thr	gtc Val	tat Tyr 150	gcg Ala	ctg Leu	atc Ile	ctg Leu	ggg Gly 155	gat Asp	cgc Arg	tgg Trp	atg Met	tac Tyr 160		480
gtg Val	gtc Val	ttc Phe	tgg Trp	ccg Pro 165	ttg Leu	ccg Pro	tcg Ser	atc Ile	ctg Leu 170	gcg Ala	tcg Ser	atc Ile	cag Gln	ctg Leu 175	ttc Phe		528
gtg Val	ttc Phe	ggc Gly	act Thr 180	Trp	ctg Leu	ccg Pro	cac His	cgc Arg 185	ccc Pro	ggc Gly	cac His	gac Asp	gcg Ala 190	Phe	ccg Pro		576
gac Asp	cgc Arg	cat His 195	Asn	gcg Ala	cgg Arg	tcg Ser	tcg ser 200	Arg	atc Ile	agc Ser	gac Asp	cct Pro 205	val	tcg Ser	ctg Leu		624
ctg Leu	acc Thr 210	Cys	ttt Phe	cat His	ttt Phe	ggc Gly 215	ĠIJ	tat Tyr	His	His	G1u 220	His	cac His	ctg Leu	cac His		672
									sei	te 2	. g						

ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp 225 230 235 240 720 729

acc gca tga Thr Ala

<210> 24

<211> 242

<212> **PRT** 

<213> Paracoccus marcusii

<400> 24

et Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu 5 10 15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His 20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala 50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn 65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp 85 90 95

Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr 100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala 115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro 130 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr 145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe 165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro 180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu 195 200 205 Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His 210 220 Pro Thr Val Pro Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp 225 230 235 Thr Ala <210> 25 <211> 1629 112> DNA 13> Synechococystis <220> <221> **CDS** <222> (1)..(1629)<223> <400> 25 atg atc acc acc gat gtt gtc att att ggg gcg ggg cac aat ggc tta Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu 1 5 10 15 48 gtc tgt gca gcc tat ttg ctc caa cgg ggc ttg ggg gtg acg tta cta Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu 20 25 30 96 a aag cgg gaa gta cca ggg ggg gcg gcc acc aca gaa gct ctc atg Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met 35 . 40 . 45 144 ccg gag cta tcc ccc cag ttt cgc ttt aac cgc tgt gcc att gac cac
Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His
50 55 60 192 gaa ttt atc ttt ctg ggg ccg gtg ttg cag gag cta aat tta gcc cag Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln 65 70 75 80 240 tat ggt ttg gaa tat tta ttt tgt gac ccc agt gtt ttt tgt ccg ggg Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly 85 90 95 288 ctg gat ggc caa gct ttt atg agc tac cgt tcc cta gaa aaa acc tgt Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys 100 105 110336

gcc cac att gcc acc tat agc ccc cga gat gcg gaa aaa tat cgg caa

Seite 30

04sequ.txt Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln
115 120 125 ttt gtc aat tat tgg acg gat ttg ctc aac gct gtc cag cct gct ttt Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe 130 140 432 aat gct ccg ccc cag gct tta cta gat tta gcc ctg aac tat ggt tgg Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp 150 155 160 480 gaa aac tta aaa tcc gtg ctg gcg atc gcc ggg tcg aaa acc aag gcg Glu Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala 165 170 175 528 ttg gat ttt atc cgc act atg atc ggc tcc ccg gaa gat gtg ctc aat Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn 180 185 576 gaa tgg ttc gac agc gaa cgg gtt aaa gct cct tta gct aga cta tgt Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys 195 200 205 624 g gaa att ggc gct ccc cca tcc caa aag ggt agt agc tcc ggc atg r Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met 210 220 672 atg atg gtg gcc atg cgg cat ttg gag gga att gcc aga cca aaa gga Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly 235 230 240 720 ggc act gga gcc ctc aca gaa gcc ttg gtg aag tta gtg caa gcc caa Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln 245 250 255 768 ggg gga aaa atc ctc act gac caa acc gtc aaa cgg gta ttg gtg gaa Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu 260 265 270 816 aac aac cag gcg atc ggg gtg gag gta gct aac gga gaa cag tac cgg Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg 275 280 285 864 gcc aaa aaa ggc gtg att tct aac atc gat gcc cgc cgt tta ttt ttg Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu 290 295 300 912 a ttg gtg gaa ccg ggg gcc cta gcc aag gtg aat caa aac cta ggg Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly 305 310 315 960 gaa cga ctg gaa cgg cgc act gtg aac aat aac gaa gcc att tta aaa Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys 325 330 335 1008 atc gat tgt gcc ctc tcc ggt tta ccc cac ttc act gcc atg gcc ggg Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly 340 345 350 1056 ccg gag gat cta acg gga act att ttg att gcc gac tcg gta cgc cat Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His 355 360 365 **1104** · gtc gag gaa gcc cac gcc ctc att gcc ttg ggg caa att ccc gat gct Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala 370 375 380 1152 aat ccg tct tta tat ttg gat att ccc act gta ttg gac ccc acc atg 1200 Seite 31

04sequ.txt Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met 385 390 395 400 gcc ccc cct ggg cag cac acc ctc tgg atc gaa ttt ttt gcc ccc tac Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr 405 410 415 1248 cgc atc gcc ggg ttg gaa ggg aca ggg tta atg ggc aca ggt tgg acc Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr 420 425 430 1296 gat gag tta aag gaa aaa gtg gcg gat cgg gtg att gat aaa tta acg Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr 435 440 445 1344 gac tat gcc cct aac cta aaa tct ctg atc att ggt cgc cga gtg gaa Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu 450 460 1392 agt ccc gcc gaa ctg gcc caa cgg ctg gga agt tac aac ggc aat gtc Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val 465 470 475 480 1440 t cat ctg gat atg agt ttg gac caa atg atg ttc ctc cgg cct cta r His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu 1488 490 ccg gaa att gcc aac tac caa acc ccc atc aaa aat ctt tac tta aca Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr 500 505 1536 ggg gcg ggt acc cat ccc ggt ggc tcc ata tca ggt atg ccc ggt aga Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg 515 1584 aat tgc gct cgg gtc ttt tta aaa caa caa cgt cgt ttt tgg taa Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp 530 535 540 1629

<210> 26

<211> 542

<212> PRT

-13> Synechococystis

<400> 26

Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu 10 15

Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu 20 25 30

Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met 35 40 45

Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His 50 55 60

Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln 65 70 75 80 Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys 100 105 110Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln 115 125 Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe 130 135 140 Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp 145 150 155 160 u Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala 165 170 175 Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn 180 185 190 Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys 195 200 205 Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Gly Met 210 215 220 Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly 225 230 235 240 Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln 245 250 255 Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu 260 265 270 Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg 275 280 285 Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu 290 295 300 Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly 305 310 315 320 Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys 325 330 335 The Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly 350 Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala 370 Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met 385 Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr

sp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr 435 440 445

Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu 450 460

Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val 465 470 475 480

Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu 485 490 495

Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr 500 505 510

Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg 515 520 525

Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp 530 535 540

<210> 27

· <211> 776

<212> DNA

<213> Bradyrhizobium sp.

<220>

<221> CDS

<222> (1)..(774)

<400	)> 2	27														
atg Met 1	cat His	gca Ala	gca Ala	acc Thr 5	gcc Ala	aag Lys	gct Ala	act Thr	gag Glu 10	ttc Phe	ggg	gcc Ala	tct Ser	cgg Arg 15	cgc Arg	48
gac Asp	gat Asp	gcg Ala	agg Arg 20	cag Gln	cgc Arg	cgc Arg	gtc Val	ggt Gly 25	ctc Leu	acg Thr	ctg Leu	gcc Ala	gcg Ala 30	gtc Val	atc Ile	96
atc Ile	gcc Ala	gcc Ala 35	tgg Trp	ctg Leu	gtg Val	ctg Leu	cat His 40	gtc Val	ggt Gly	ctg Leu	atg Met	ttc Phe 45	ttc Phe	tgg Trp	ccg Pro	144
ctg Leu	acc Thr 50	ctt Leu	cac His	agc Ser	ctg Leu	ctg Leu 55	ccg Pro	gct Ala	ttg Leu	cct Pro	ctg Leu 60	gtg Val	gtg Val	ctg Leu	cag Gln	192
acc Thr	tgg Trp	ctc Leu	tat Tyr	gta Val	ggc Gly 70	ctg Leu	ttc Phe	atc Ile	atc Ile	gcg Ala 75	cat His	gac Asp	tgc Cys	atg Met	cac His 80	240
ggc Gly	tcg Ser	ctg Leu	gtg Val	ccg Pro 85	ttc Phe	aag Lys	ccg Pro	cag Gln	gtc Val 90	aac Asn	cgc Arg	cgt Arg	atc Ile	gga Gly 95	cag Gln	288
ctc Leu	tgc Cys	ctg Leu	ttc Phe 100	ctc Leu	tat Tyr	gcc Ala	ggg Gly	ttc Phe 105	tcc Ser	ttc Phe	gac Asp	gct Ala	ctc Leu 110	aat Asn	gtc Val	336
gag Glu	cac His	cac His 115	aag Lys	cat His	cac His	cgc Arg	cat His 120	ccc Pro	ggc Gly	acg Thr	gcc Ala	gag Glu 125	gat Asp	ccc Pro	gat Asp	384
ttc Phe	gac Asp 130	gag Glu	gtg Val	ccg Pro	ccg Pro	cac His 135	ggc Gly	ttc Phe	tgg Trp	cac His	tgg Trp 140	ttc Phe	gcc Ala	agc Ser	ttt Phe	432
ttc .Phe 145	ctg Leu	cac His	tat Tyr	ttc Phe	ggc Gly 150	tgg Trp	aag Lys	cag Gln	gtc Val	gcg Ala 155	atc Ile	atc Ile	gca Ala	gcc Ala	gtc Val 160	480
g	ctg Leu	gtt Val	tat Tyr	cag Gln 165	ctc Leu	gtc Val	ttc Phe	gcc Ala	gtt Val 170	ccc Pro	ttg Leu	cag Gln	aac Asn	atc Ile 175	ctg Leu	528
ctg Leu	ttc Phe	tgg Trp	gcg Ala 180	Leu	ccc Pro	ggg Gly	ctg Leu	ctg Leu 185	ser	gcg Ala	ctg Leu	cag Gln	ctg Leu 190	Phe	acc Thr	576
ttc Phe	ggc Gly	acc Thr 195	tat Tyr	ctg Leu	ccg Pro	cac His	aag Lys 200	ccg Pro	gcc Ala	acg Thr	cag Gln	ccc Pro 205	ttc Phe	gcc Ala	gat Asp	624
cgc Arg	cac His 210	aac Asn	gcg Ala	cgg Arg	acg Thr	agc Ser 215	gaa Glu	ttt Phe	ccc Pro	gcg Ala	tgg Trp 220	ctg Leu	tcg Ser	ctg Leu	ctg Leu	672
acc Thr 225	tgc Cys	ttc Phe	cac His	ttc Phe	ggc Gly 230	ttt Phe	cat His	cac His	gag Glu	cat His 235	cat His	ctg Leu	cat His	ccc Pro	gat Asp 240	720
gcg Ala	ccg Pro	tgg Trp	tgg Trp	cgg Arg 245	ctg Leu	ccg Pro	gag Glu	atc Ile	Lys 250	cgg Arg te 3	cgg Arg	gcc Ala	ctg Leu	gaa Glu 255	Arg	768

cgt gac ta Arg Asp

<210> 28

<211> 258

<21:2> PRT

<213> Bradyrhizobium sp.

<400> 28

Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg 1 5 10 15

sp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile 20 25 30

Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro 35 40 45

Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln 50 60

Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His 65 70 75 80

Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln 85 90 95

Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val 100 105 110

u His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp 115 120 125

Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe 130 135 140

Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val 145 150 160

Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu 165 170 175

Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr 180 185 190

Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp 195 200 205

Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu 210 215 220 Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp 225 230 235 240 Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg 245 250 255 Arg Asp <210> 29 <211> 777 212> DNA 213> Nostoc sp. <220> <221> **CDS** <222> (1)..(777)<223> <400> 29 atg gtt cag tgt caa cca tca tct ctg cat tca gaa aaa ctg gtg tta Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu 1 10 15 48 ttg tca tcg aca atc aga gat gat aaa aat att aat aag ggt ata ttt Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe 20 25 30 96 It gcc tgc ttt atc tta ttt tta tgg gca att agt tta atc tta tta Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu 35 40 144 ctc tca ata gat aca tcc ata att cat aag agc tta tta ggt ata gcc Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala 50 55 60 192 atg ctt tgg cag acc ttc tta tat aca ggt tta ttt att act gct cat Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His 65 70 75 80 240 gat gcc atg cac ggc gta gtt tat ccc aaa aat ccc aga ata aat aat Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn 85 90 95 288 ttt ata ggt aag ctc act cta atc ttg tat gga cta ctc cct tat aaa Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys 100 105 110 336 105 gat tta ttg aaa aaa cat tgg tta cac cac gga cat cct ggt act gat 384

Seite 37

04sequ.txt Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp tta gac cct gat tat tac aat ggt cat ccc caa aac ttc ttt ctt tgg Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp 130 135 140 432 tat cta cat ttt atg aag tct tat tgg cga tgg acg caa att ttc gga Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly 145 150 160 480 tta gtg atg att ttt cat gga ctt aaa aat ctg gtg cat ata cca gaa Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu 528 aat aat tta att ata ttt tgg atg ata cct tct att tta agt tca gta Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val 576 180 caa cta ttt tat ttt ggt aca ttt ttg cct cat aaa aag cta gaa ggt Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly 195 200 205 624 It tat act aac ccc cat tgt gcg cgc agt atc cca tta cct ctt ttt Bly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe 210 220 672 tgg tct ttt gtt act tgt tat cac ttc ggc tac cac aag gaa cat cac Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His 225 230 235 240 720 gaa tac cct caa ctt cct tgg tgg aaa tta cct gaa gct cac aaa ata Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile 245 250. 768 777 tct tta taa Ser Leu

<210> 30

<211> 258

<212> PRT

13> Nostoc sp.

<400> 30

Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu 1 10 15

Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe 20 25 30

Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu 35 40 45

Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala 50 . 60

Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His 80 Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys 110 Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp 125 Gly Thr Asp 130 Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp 130 Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp 145 Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly 160 Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe 2215

Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile 245 250 255

Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His 225 230 235 240

Ser Leu

<210> 31

<211> 831

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (1)..(831)

<400 atg Met 1	cca	1 tcc Ser	gag Glu	tcg ser 5	tca Ser	gac Asp	gca Ala	gct Ala	cgt Arg 10	cct Pro	gtg Val	ttg Leu	aag Lys	cac His 15	gcc Ala		48
tat Tyr	aaa Lys	cct Pro	cca Pro 20	gca Ala	tct Ser	gac Asp	gcc Ala	aag Lys 25	ggc Gly	atc Ile	act Thr	atg Met	gcg Ala 30	ctg Leu	acc Thr		96
atc Ile	att Ile	ggc Gly 35	acc Thr	tgg Trp	acc Thr	gca Ala	gtg Va 1 40	ttt Phe	tta Leu	cac His	gca Ala	ata Ile 45	ttc Phe	caa Gln	atc Ile	:	144
agg Arg	cta Leu 50	ccg Pro	aca Thr	tcc Ser	atg Met	gac Asp 55	cag Gln	ctt Leu	cac His	tgg Trp	ttg Leu 60	cct Pro	gtg Val	tcc Ser	gaa Glu	;	192
gcc 11a	aca Thr	gcc Ala	cag Gln	ctg Leu	ttg Leu 70	ggc Gly	gga Gly	agc Ser	agc Ser	agc Ser 75	cta Leu	ttg Leu	cac His	atc Ile	gcc Ala 80		240
gca Ala	gtc Val	ttc Phe	att Ile	gta Val 85	ctt Leu	gag Glu	ttt Phe	ctg Leu	tac Tyr 90	act Thr	ggt Gly	cta Leu	ttc Phe	atc Ile 95	acc Thr		288
acg Thr	cat His	gat Asp	gca Ala 100	Met	cat His	ggc Gly	acc Thr	ata Ile 105	gct Ala	ttg Leu	agg Arg	aac Asn	agg Arg 110	Gln	ctc Leu		336
aat Asn	gat Asp	ctc Leu 115	Leu	ggc Gly	aac Asn	atc Ile	tgc Cys 120	Ile	tca Ser	ctg Leu	tac Tyr	gcc Ala 125	Trp	ttt Phe	gac Asp		384
tac Tyr	agc Ser 130	Met	cac His	tgg Trp	gag Glu	cac His 135	His	aac Asn	cat His	act Thr	ggc Gly 140	Glu	gtg Val	ggg Gly	aaa Lys		432
gac Asp 145	Pro	gac Asp	tto Phe	cac His	aaa Lys 150	Gly	aat Asn	cct Pro	ggc	ctt Leu 155	ı Val	Pro	tgg Trp	tto Phe	gcc Ala 160		480
	ttc Phe	atg Met	tco Ser	agc Ser 165	Tyr	atg Met	tcc Ser	ctg Leu	tgg Trp 170	Glr	ttt Phe	gco Ala	cgg Arg	cto Lei 17!	g gca u Ala		528
tgg Trp	tgg Trp	gca Ala	gtg Val 180	Val	atg Met	caa Glr	acg Thr	ttg Leu 185	ı Giy	g gco / Ala	c cco a Pro	ato Me	g gcg t Ala 190	a ASI	t ctc 1 Leu		576
cta Lei	gto Val	tto Phe 195	e Met	g gct Ala	gca Ala	gco a Ala	c cca R Pro 200	) Ile	ttg Lei	g tca u Sei	a gca r Ala	a tto a Pho 20	e Ar	c cto g Leo	c ttc u Phe		624
tac Tyr	tto Phe 210	e Gly	c act	t tac r Tyr	cto Lei	y cca u Pro 21!	His	c aag s Lys	g cci	t gag o Gli	g cca u Pro 22	o Gi	y Pr	t gc o Al	a gca a Ala		672
gg( G1) 22:	/ Sei	caq Gli	g gte n Va	c ato 1 Met	5 tc1 5 Sei 230	r Tri	g tto Pho	c agg e Arg	g gc g Ala	c aa a Ly: 23	<u>s</u> Th	a ag r Se	t ga r Gl	g gc u Al	a tct a Ser 240		720
ga1 As <b>j</b>	t gtg o Va	g atg	g age t Se	c tto r Phe 245	e Lei	g aca u Th	a tge r Cy:	c tad	r Hi 25	<u>s</u> Ph	e As	c ct p Le	g tt u Ph	t gc e Al 25	c ccc a Pro 5		768

tgg tgg cag ctg ccc cac tgc cgc cgc ctg tct ggg cgt ggc ctg gtg Trp Trp Gln Leu Pro His Cys Arg Arg Leu Ser Gly Arg Gly Leu Val 260 265 270

816

cct gcc ttg gca tga Pro Ala Leu Ala 275

831

<210> 32

<211> 276

<212> PRT

<213> Haematococcus pluvialis

<400> 32

et Pro Ser Glu Ser Ser Asp Ala Ala Arg Pro Val Leu Lys His Ala 5 10 15

Tyr Lys Pro Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr 20 25 30

Ile Ile Gly Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile 35 40 45

Arg Leu Pro Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu 50 60

Ala Thr Ala Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala 65 70 75 80

Ala Val Phe Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr 85 90 95

r His Asp Ala Met His Gly Thr Ile Ala Leu Arg Asn Arg Gln Leu 100 105 110

Asn Asp Leu Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp 115 120 125

Tyr Ser Met His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys 130 135 140

Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val Pro Trp Phe Ala 145 150 155 160

Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe Ala Arg Leu Ala 165 170 175

Trp Trp Ala Val Val Met Gln Thr Leu Gly Ala Pro Met Ala Asn Leu 180 185 190 Seite 41

Leu Val Phe 195	Met A	la Ala	Ala	Pro 200	Ile	Leu	Ser	Ala	Phe 205	Arg	Leu	Phe
Tyr Phe Gly 210	Thr T	yr Leu	Pro 215	ніѕ	Lys	Pro	Glu	Pro 220	Glу	Pro	Ala	Ala
Gly Ser Gln 225	Val M	et Ser 230	Тгр	Phe	Arg	Ala	Lys 235	Thr	Ser	Glu	Ala	Ser 240
Asp Val Met	Ser P	he Leu 45	Thr	Cys	Tyr	ніs 250	Phe	Asp	Leu	Phe	Ala 255	Pro
Trp Trp Gln	Leu P 260	ro His	Cys	Arg	Arg 265	Leu	Ser	Gly	Arg	Gly 270	Leu	Val

ero Ala Leu Ala 275

<210> 33

<211> 729

<212> DNA

<213> Paracoccus sp. MBIC1143

<220>

<221> CDS

<222> (1)..(729)

<223>

.2.	00> g agc t Ser	33 gca Ala	cat His	gcc Ala 5	ctg Leu	ccc Pro	aag Lys	gca Ala	gat Asp 10	ctg Leu	acc Thr	gcc Ala	acc Thr	agc Ser 15	ctg Leu	48
at Il	c gtc e Val	tcg Ser	ggc Gly 20	ggc Gly	atc Ile	atc Ile	gcc Ala	gct Ala 25	tgg Trp	ctg Leu	gcc Ala	ctg Leu	cat His 30	gtg Val	cat His	96
gc Al	g cto a Lei	tgg Trp 35	ttt Phe	ctg Leu	gac Asp	gca Ala	gcg Ala 40	gcg Ala	cat His	ccc Pro	atc Ile	ctg Leu 45	gcg Ala	atc Ile	gca Ala	. 144
aa As	t tto n Phe 50	ctg Leu	ggg Gly	ctg Leu	acc Thr	tgg Trp 55	ctg Leu	tcg Ser	gtc Val	gga Gly	ttg Leu 60	ttc Phe	atc Ile	atc Ile	gcg Ala	192
са Ні 65	t gad s Asp	gcg Ala	atg Met	cac His	ggg Gly 70	tcg Ser	gtg Val	gtg Val	ccg Pro	ggg Gly 75	cgt Arg	ccg Pro	cgc Arg	gcc Ala	aat Asn 80	240
gc	g gc	g atg	ggc	cag	ctt	gtc	ctg	tgg	_	tat te 4	_	gga	ttt	tcg	tgg	288

								(	)4sed	u.tx	(t						
Ala	Ala	Met	Gly	Gln 85	Leu	۷a٦	Leu	Trp	Leu 90	Tyr	Ala	Gly	Phe	Ser 95	Trp		
cgc Arg	aag Lys	atg Met	atc Ile 100	gtc Val	aag Lys	cac His	atg Met	gcc Ala 105	cat His	cac His	cgc Arg	cat His	gcc Ala 110	gga Gly	acc Thr	3	336
gac Asp	gac Asp	gac Asp 115	ccc Pro	gat Asp	ttc Phe	gac Asp	cat His 120	ggc Gly	ggc Gly	ccg Pro	gtc Val	cgc Arg 125	tgg Trp	tac Tyr	gcc Ala	3	384
cgc Arg	ttc Phe 130	atc Ile	ggc Gly	acc Thr	tat Tyr	ttc Phe 135	ggc Gly	tgg Trp	cgc Arg	gag Glu	ggg Gly 140	ctg Leu	ctg Leu	ctg Leu	CCC Pro	4	432
gtc Val 145	atc Ile	gtg Val	acg Thr	gtc Val	tat Tyr 150	gcg Ala	ctg Leu	atc Ile	ctt Leu	ggg Gly 155	gat Asp	cgc Arg	tgg Trp	atg Met	tac Tyr 160	4	480
gtg Val	gtc val	ttc Phe	tgg Trp	ccg Pro 165	ctg Leu	ccg Pro	tcg Ser	atc Ile	ctg Leu 170	gcg Ala	tcg Ser	atc Ile	cag Gln	ctg Leu 175	ttc Phe	•	528
ig il	ttc Phe	ggc Gly	acc Thr 180	tgg Trp	ctg Leu	ccg Pro	cac His	cgc Arg 185	ccc Pro	ggc Gly	cac His	gac Asp	gcg Ala 190	ttc Phe	ccg Pro		576
gac Asp	cgc Arg	cac His 195	aat Asn	gcg Ala	cgg Arg	tcg Ser	tcg Ser 200	cgg Arg	atc Ile	agc Ser	gac Asp	ccc Pro 205	gtg Val	tcg Ser	ctg Leu	(	624
ctg Leu	acc Thr 210	tgc Cys	ttt Phe	cac His	ttt Phe	ggc Gly 215	ggt Gly	tat Tyr	cat His	cac His	gaa Glu 220	cac His	cac His	ctg Leu	cac His	1	672
ccg Pro 225	acg Thr	gtg Val	ccg Pro	tgg Trp	tgg Trp 230	cgc Arg	ctg Leu	ccc Pro	agc Ser	acc Thr 235	cgc Arg	acc Thr	aag Lys	ggg Gly	gac Asp 240		720
	gca Ala	tga															729

<210> 34

11> 242

<212> PRT

<213> Paracoccus sp. MBIC1143

<400> 34

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu  $10 \hspace{1cm} 15$ 

·Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His 20 25 30

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
90 Tyr Ala Gly Phe Ser Trp
Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr
Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
11 Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 Ile Val Thr Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His 210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp 230 235 . 240

Thr Ala

<210> 35

<211> 735

<212> DNA

<213> Brevundimonas aurantiaca

<220>

<221> CDS

<222> (1)..(735)

<400 atg			acc	atc	מככ	nan	cca	כמכ	acc	atc	cca	cac	cad	acc	taa	48
Met 1	Thr	Ala	Ala	val 5	Ala	ฐาน ฐาน	Pro	Arg	Thr 10	val	Pro	Arg	ĞÎñ	Thr 15	Trp	
atc Ile	ggt Gly	ctg Leu	acc Thr 20	ctg Leu	gcg Ala	gga Gly	atg Met	atc Ile 25	gtg Val	gcg Ala	gga Gly	tgg Trp	gcg Ala 30	gtt Val	ctg Leu	96
cat His	gtc Val	tac Tyr 35	ggc Gly	gtc Val	tat Tyr	ttt Phe	cac His 40	cga Arg	tgg Trp	ggg Gly	ccg Pro	ttg Leu 45	acc Thr	ctg Leu	gtg Val	144
atc Ile	gcc Ala 50	ccg Pro	gcg Ala	atc Ile	gtg Val	gcg Ala 55	gtc Val	cag Gln	acc Thr	tgg Trp	ttg Leu 60	tcg Ser	gtc Val	ggc Gly	ctt Leu	192
ttc	atc Ile	gtc Val	gcc Ala	cat His	gac Asp 70	gcc Ala	atg Met	tac Tyr	ggc Gly	tcc Ser 75	ctg Leu	gcg Ala	ccg Pro	gga Gly	cgg Arg 80	240
ccg Pro	cgg Arg	ctg Leu	aac Asn	gcc Ala 85	gca Ala	gtc val	ggc Gly	cgg Arg	ctg Leu 90	acc Thr	ctg Leu	ggg Gly	ctc Leu	tat Tyr 95	gcg Ala	288
ggc Gly	ttc Phe	cgc Arg	ttc Phe 100	Asp	cgg Arg	ctg Leu	aag Lys	acg Thr 105	gcg Ala	cac His	cac His	gcc Ala	cac His 110	cac His	gcc Ala	336
gcg Ala	CCC Pro	ggc Gly 115	acg Thr	gcc Ala	gac Asp	gac Asp	ccg Pro 120	gat Asp	ttt Phe	cac His	gcc Ala	ccg Pro 125	gcg Ala	ccc Pro	cgc Arg	384
gcc Ala	ttc Phe 130	Leu	ccc Pro	tgg Trp	ttc Phe	ctg Leu 135	Asn	ttc Phe	ttt Phe	cgc Arg	acc Thr 140	Tyr	ttc Phe	ggc Gly	tgg Trp	432
cgc Arg 145	gag Glu	atg Met	gcg Ala	gtc Val	ctg Leu 150	Thr	gcc Ala	ctg Leu	gtc Val	ctg Leu 155	Ile	gcc Ala	ctc Leu	ttc Phe	ggc Gly 160	480
g u	ggg Gly	gcg Ala	cgg Arg	ccg Pro 165	Ala	aat Asn	ctc Leu	ctg Leu	acc Thr 170	Phe	tgg Trp	gcc Ala	gcg Ala	ccg Pro 175	gcc Ala	528
ctg Leu	ctt Leu	tca Ser	gcg Ala 180	Leu	cag Gln	cto Leu	tto Phe	acc Thr 185	· Phe	ggc	acc Thr	tgg Trp	ctg Leu 190	Pro	cac His	576
cgc Arg	cac His	acc Thr 195	Asp	cag Glr	ccg Pro	tto Phe	gcc Ala 200	LASP	gcg Ala	cac His	cac His	gco Ala 205	ı Arg	ago g Ser	agc Ser	624
ggc Gly	tac Tyr 210	· Gly	cco Pro	gto Val	ctt Lei	tco Ser 215	Lei	cto Lei	acc Thr	tgt Cys	t tto s Phe 220	e His	tto Phe	e Gly	cgc Arg	672
cac His 225	: His	gaa Glu	a cad u His	cat His	cto Lei 230	ı Sei	cco Pro	tgg Tr	g cgg o Arg	p cco p Pro 235	Trp	g tgg o Tri	g cgt o Arg	t cto	g tgg u Trp 240	720
cgc Arg	gg Gly	gag Gli	j tci i Sei	t tga	ı											735

<210> 36

<211> 244

<212> PRT

<213> Brevundimonas aurantiaca

<400> 36

Met Thr Ala Ala Val Ala Glu Pro Arg Thr Val Pro Arg Gln Thr Trp 5 10 15

Ile Gly Leu Thr Leu Ala Gly Met Ile Val Ala Gly Trp Ala Val Leu 20 25 30

His Val Tyr Gly Val Tyr Phe His Arg Trp Gly Pro Leu Thr Leu Val 35 40 45

Ile Ala Pro Ala Ile Val Ala Val Gln Thr Trp Leu Ser Val Gly Leu 50 60

Phe Ile Val Ala His Asp Ala Met Tyr Gly Ser Leu Ala Pro Gly Arg 65 70 75 80

Pro Arg Leu Asn Ala Ala Val Gly Arg Leu Thr Leu Gly Leu Tyr Ala 85 90 95

Gly Phe Arg Phe Asp Arg Leu Lys Thr Ala His His Ala His Ala 100 105 110

Ala Pro Gly Thr Ala Asp Asp Pro Asp Phe His Ala Pro Ala Pro Arg 115 120 125

a Phe Leu Pro Trp Phe Leu Asn Phe Phe Arg Thr Tyr Phe Gly Trp 130 140

Arg Glu Met Ala Val Leu Thr Ala Leu Val Leu Ile Ala Leu Phe Gly 145 150 155 160

Leu Gly Ala Arg Pro Ala Asn Leu Leu Thr Phe Trp Ala Ala Pro Ala 165 170 175

Leu Leu Ser Ala Leu Gln Leu Phe Thr Phe Gly Thr Trp Leu Pro His 180 185 190

Arg His Thr Asp Gln Pro Phe Ala Asp Ala His His Ala Arg Ser Ser 195 200 205

Gly Tyr Gly Pro Val Leu Ser Leu Leu Thr Cys Phe His Phe Gly Arg 210 215 220 Seite 46

His His Glu His His	Leu Ser Pro	Trp Arg Pro	Trp Trp Arg	Leu Trp
225	230	235		240

Arg Gly Glu Ser

<210> 37

<211> 690

<212> DNA

<213> Nodularia spumigena NSOR10

<220>

221> CDS

222> (1)..(690)

<223>

<400	> 3	7													++-	48	2
atg Met 1	gcg Ala	atc Ile	gcc Ala	att Ile 5	att Ile	agt Ser	ata Ile	ırp	gct Ala 10	atc Ile	agc Ser	<b>Leu</b>	Gly	Leu 15	Leu	40	,
ctt Leu	tat Tyr	att Ile	gat Asp 20	ata Ile	tcc Ser	caa Gln	ttc Phe	aag Lys 25	ttt Phe	tgg Trp	atg Met	ttg Leu	tta Leu 30	ccg Pro	ctc Leu	90	5
ata Ile	ttt Phe	tgg Trp 35	caa Gln	aca Thr	ttt Phe	tta Leu	tat Tyr 40	acg Thr	gga Gly	tta Leu	ttt Phe	att Ile 45	aca Thr	gct Ala	cat His	14	4
gat Asp	gcc Ala 50	atg Met	cat His	ggg Gly	gta Val	gtt Val 55	ttt Phe	ccc Pro	aaa Lys	aat Asn	ccc Pro 60	aaa Lys	atc Ile	aac Asn	cat His	19	2
Phe 65	att Ile	ggc Gly	tca Ser	ttg Leu	tgc Cys 70	ctg Leu	ttt Phe	ctt Leu	tat Tyr	ggt Gly 75	ctt Leu	tta Leu	cct Pro	tat Tyr	caa Gln 80	24	0
aaa Lys	ctt Leu	tta Leu	aaa Lys	aag Lys 85	cat His	tgg Trp	cta Leu	cat His	cac His 90	cat His	aat Asn	cca Pro	gcc	agt Ser 95	gaa Glu	28	8
aca Thr	gat Asp	cca Pro	gat Asp 100	Pne	cac His	aac Asn	ggg Gly	aag Lys 105	GII	aaa Lys	aac Asn	ttt Phe	ttt Phe 110		tgg Trp,		36
tat Tyr	tta Leu	tat Tyr 115	Phe	ato Met	aag Lys	g cgt s Arg	tac Tyr 120	. irb	agt Ser	tgg Trp	tta Lei	caa Glr 125	,	ato Ile	aca Thr	. 38	34
tta Leu	ato Met	: Ile	ati	t tai	t aad r Asi	tta Lei 13	ı Lei	ı aaa ı Lys	tai Tyi	t ata r Ile	a tgg E Trp 140	יום כ	t tti s Phe	cca e Pro	a gag o Glu	4:	32
gat	: aa1	t atg	g ac	t ta	t tt	t tg	g gta	a gtt	se se	c tca	a at <sup>.</sup> 47	t tt	a ag	t tc	t tta	4	80

O4sequ.txt Asp Asn Met Thr Tyr Phe Trp Val Val Pro Ser Ile Leu Ser Ser Leu	
145 150 155 160	
caa tta ttt tat ttt gga act ttt cta ccc cac agt gag cct gta gaa Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Ser Glu Pro Val Glu 165 170 175	528
ggt tat aaa gag cct cat cgt tcc caa act att agc cgt ccc att tgg Gly Tyr Lys Glu Pro His Arg Ser Gln Thr Ile Ser Arg Pro Ile Trp 180 185 190	576
tgg tca ttt ata act tgt tac cat ttt ggt tat cat tac gaa cat cat Trp Ser Phe Ile Thr Cys Tyr His Phe Gly Tyr His Tyr Glu His His 195 200 205	624
gaa tac ccc cat gtt cct tgg tgg caa tta cca gaa att tat aaa atg Glu Tyr Pro His Val Pro Trp Trp Gln Leu Pro Glu Ile Tyr Lys Met 210 215 220	672
tct aaa tca aat ttg tga Ser Lys Ser Asn Leu 225	690
£10> 38	
<211> 229	
<212> PRT	
<213> Nodularia spumigena NSOR10	
<400> 38	
<pre>&lt;400&gt; 38 Met Ala Ile Ala Ile Ile Ser Ile Trp Ala Ile Ser Leu Gly Leu Leu 1 5 10 15</pre>	
Met Ala Ile Ala Ile Ile Ser Ile Trp Ala Ile Ser Leu Gly Leu	
Met Ala Ile Ala Ile Ile Ser Ile Trp Ala Ile Ser Leu Gly Leu Leu  1 10 15  Leu Tyr Ile Asp Ile Ser Gln Phe Lys Phe Trp Met Leu Leu Pro Leu	
Met Ala Ile Ala Ile Ile Ser Ile Trp Ala Ile Ser Leu Gly Leu Leu 15  Leu Tyr Ile Asp Ile Ser Gln Phe Lys Phe Trp Met Leu Leu Pro Leu 20  Tle Phe Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His	
Met Ala Ile Ala Ile Ile Ser Ile Trp Ala Ile Ser Leu Gly Leu Leu 15  Leu Tyr Ile Asp Ile Ser Gln Phe Lys Phe Trp Met Leu Leu Pro Leu 20  Ile Phe Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His 45  Asp Ala Met His Gly Val Val Phe Pro Lys Asn Pro Lys Ile Asn His	
Met Ala Ile Ala Ile Ile Ser Ile Trp Ala Ile Ser Leu Gly Leu Leu Leu Tyr Ile Asp Ile Ser Gln Phe Lys Phe Trp Met Leu Leu Pro Leu 20 Ile Phe Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His 45 Asp Ala Met His Gly Val Val Phe Pro Lys Asn Pro Lys Ile Asn His 60 Phe Ile Gly Ser Leu Cys Leu Phe Leu Tyr Gly Leu Leu Pro Tyr Gln	
Met Ala Ile Ala Ile Ile Ser Ile Trp Ala Ile Ser Leu Gly Leu Leu 15  Leu Tyr Ile Asp Ile Ser Gln Phe Lys Phe Trp Met Leu Leu Pro Leu 20  Tle Phe Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His 45  Asp Ala Met His Gly Val Val Phe Pro Lys Asn Pro Lys Ile Asn His 50  Phe Ile Gly Ser Leu Cys Leu Phe Leu Tyr Gly Leu Leu Pro Tyr Gln 80  Lys Leu Leu Lys Lys His Trp Leu His His His Asn Pro Ala Ser Glu	

04sequ.txt Leu Met Ile Ile Tyr Asn Leu Leu Lys Tyr Ile Trp His Phe Pro Glu 135 Asp Asn Met Thr Tyr Phe Trp Val Val Pro Ser Ile Leu Ser Ser Leu 145 150 155 160 Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Ser Glu Pro Val Glu Gly Tyr Lys Glu Pro His Arg Ser Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile Thr Cys Tyr His Phe Gly Tyr His Tyr Glu His His 195 200 205 Glu Tyr Pro His Val Pro Trp Trp Gln Leu Pro Glu Ile Tyr Lys Met 210 215 220 Lys Ser Asn Leu <210> 789 <211> <212> DNA Nostoc punctiforme ATCC 29133 <213> <220> <221> CDS (1)..(789)<222> ≤223> <400> ttg aat ttt tgt gat aaa cca gtt agc tat tat gtt gca ata gag caa Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln 1 5 10 48 tta agt gct aaa gaa gat act gtt tgg ggg ctg gtg att gtc ata gta Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val 20 25 30 96 att att agt ctt tgg gta gct agt ttg gct ttt tta cta gct att aat Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn 35 40 45 144 tat gcc aaa gtc cca att tgg ttg ata cct att gca ata gtt tgg caa Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln 50 . 55 60 192 atg ttc ctt tat aca ggg cta ttt att act gca cat gat gct atg cat Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His 65 70 75 80 240

Seite 49

										,						
ggg Gly	tca Ser	gtt Val	tat Tyr	cgt Arg 85	aaa Lys	aat Asn	ccc Pro	aaa Lys	att Ile 90	aat Asn	aat Asn	ttt Phe	atc Ile	ggt Gly 95	tca Ser	288
cta Leu	gct Ala	gta Val	gcg Ala 100	ctt Leu	tac Tyr	gct Ala	gtg Val	ttt Phe 105	cca Pro	tat Tyr	caa Gln	cag Gln	atg Met 110	tta Leu	aag Lys	336
aat Asn	cat His	tgc Cys 115	tta Leu	cat His	cat His	cgt Arg	cat His 120	cct Pro	gct Ala	agc Ser	gaa Glu	gtt Val 125	gac Asp	cca Pro	gat Asp	384
ttt Phe	cat His 130	gat Asp	ggt Gly	aag Lys	aga Arg	aca Thr 135	aac Asn	gct Ala	att Ile	ttc Phe	tgg Trp 140	tat Tyr	ctc Leu	cat His	ttc Phe	432
atg Met 145	ata Ile	gaa Glu	tac Tyr	tcc Ser	agt Ser 150	tgg Trp	caa Gln	cag Gln	tta Leu	ata Ile 155	gta Val	cta Leu	act Thr	atc Ile	cta Leu 160	480
ttt Phe	aat Asn	tta Leu	gct Ala	aaa Lys 165	tac Tyr	gtt Val	ttg Leu	cac His	atc Ile 170	cat His	caa Gln	ata Ile	aat Asn	ctc Leu 175	atc Ile	528
tta Leu	ttt Phe	tgg Trp	agt Ser 180	Ile	cct Pro	cca Pro	att Ile	tta Leu 185	agt Ser	tcc Ser	att Ile	caa Gln	ctg Leu 190	Phe	tat Tyr	576
ttc Phe	gga Gly	aca Thr 195	ttt Phe	ttg Leu	cct Pro	cat His	cga Arg 200	Glu	ccc Pro	aag Lys	aaa Lys	gga Gly 205	Tyr	gtt Val	tat Tyr	624
ccc Pro	cat His 210	Cys	agc Ser	caa Gln	aca Thr	ata Ile 215	Lys	ttg Leu	cca Pro	act Thr	ttt Phe 220	Leu	tca Ser	ttt Phe	atc Ile	672
gct Ala 225	L Cys	tac Tyr	cac His	ttt Phe	ggt Gly 230	Tyr	cat His	gaa Glu	gaa Glu	cat His 235	His	gag Glu	tat Tyr	ccc Pro	cat His 240	720
gta Val	cct Pro	tgg Trp	tgg Trp	caa Gln 245	Leu	cca Pro	tct Ser	gta Val	tat Tyr 250	Lys	cag Glr	aga n Arg	gta y Val	tto Phe 255	aac Asn	768
	tca Ser			' Asn			ı									789

<210> 40

<211> 262

<212> PRT

<213> Nostoc punctiforme ATCC 29133

<400> 40

Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln 10 15

Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val 20 25 30 Seite 50

Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln

Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His 80

Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser

Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys

Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp 115 . 120 125

Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe 130 135 140

Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu 145 150 160

Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile 165 170 175

Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr 180 185 190

Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr 195 200 205

o His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile 210 215 220

Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His 225 230 235 240

Val Pro Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn 245 250 255

Asn Ser Val Thr Asn Ser 260

<210> 41

<211> 762

<212> DNA

<213> Nostoc punctiforme ATCC 29133

<220	>															
<221:	> C	DS														
<222:	> (	1)	(762	)												
<223	>															
<400 gtg val 1	atc	cag	tta Leu	gaa Glu 5	caa Gln	cca Pro	ctc Leu	agt Ser	cat His 10	caa Gln	gca Ala	aaa Lys	ctg Leu	act Thr 15	cca Pro	48
gta Val	ctg Leu	aga Arg	agt Ser 20	aaa Lys	tct Ser	cag Gln	ttt Phe	aag Lys 25	ggg Gly	ctt Leu	ttc Phe	att Ile	gct Ala 30	att Ile	gtc Val	96
tt le	gtt Val	agc Ser 35	gca Ala	tgg Trp	gtc Val	att Ile	agc Ser 40	ctg Leu	agt Ser	tta Leu	tta Leu	ctt Leu 45	tcc Ser	ctt Leu	gac Asp	144
atc Ile	tca Ser 50	aag Lys	cta Leu	aaa Lys	ttt Phe	tgg Trp 55	atg Met	tta Leu	ttg Leu	cct Pro	gtt Val 60	ata Ile	cta Leu	tgg Trp	caa Gln	192
aca Thr 65	ttt Phe	tta Leu	tat Tyr	acg Thr	gga Gly 70	tta Leu	ttt Phe	att Ile	aca Thr	tct Ser 75	cat His	gat Asp	gcc Ala	atg Met	cat His 80	240
ggc Gly	gta Val	gta Val	ttt Phe	ccc Pro 85	caa Gln	aac Asn	acc Thr	aag Lys	att Ile 90	aat Asn	cat His	ttg Leu	att Ile	gga Gly 95	aca Thr	288
ttg Leu	acc Thr	cta Leu	tcc Ser 100	ctt Leu	tat Tyr	ggt Gly	ctt Leu	tta Leu 105	cca Pro	tat Tyr	caa Gln	aaa Lys	cta Leu 110	ttg Leu	aaa Lys	336
aaa Lys	cat His	tgg Trp 115	Leu	cac His	cac His	cac His	aat Asn 120	cca Pro	gca Ala	agc Ser	tca Ser	ata Ile 125	gac Asp	ccg Pro	gat Asp	384
Phe	cac His 130	Asn	ggt Gly	aaa Lys	cac His	caa Gln 135	Ser	ttc Phe	ttt Phe	gct Ala	tgg Trp 140	tat Tyr	ttt Phe	cat His	ttt Phė	432
atg Met 145	Lys	ggt Gly	tac Tyr	tgg Trp	agt Ser 150	Trp	ggg Gly	caa Gln	ata Ile	att Ile 155	Ala	ttg Leu	act Thr	att Ile	att Ile 160	480
tat Tyr	aac Asn	ttt Phe	gct Ala	aaa Lys 165	Tyr	ata Ile	ctc Leu	cat His	atc Ile 170	Pro	agt Ser	gat Asp	aat Asn	cta Leu 175	act Thr	528
tac Tyr	ttt Phe	tgg Trp	gtg Val 180	Leu	ccc Pro	tcg Ser	ctt Leu	tta Leu 185	. Ser	tca Ser	tta Leu	caa Glr	tta Leu 190	ı Ph€	tat Tyr	576
ttt Phe	ggt	act Thr 195	Phe	tta Leu	ccc Pro	cat His	agt Ser 200	Glr	ı cca ı Pro	ata Ile	ggg Gly	ggt Gly 205	/ Tyr	gti Va	cag I Gln	624

cct cat tgt gcc caa aca att agc cgt cct att tgg tgg tca ttt atc

Seite 52

672

04sequ.txt Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile 210 215 220 acg tgc tat cat ttt ggc tac cac gag gaa cat cac gaa tat cct cat Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His 225 230 235 240 720 att tct tgg tgg cag tta cca gaa att tac aaa gca aaa tag Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys 245 250 762 <210> 42

<211> 253

<212> **PRT** 

Nostoc punctiforme ATCC 29133 <213>

<400> 42

al Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro 5 10 15

Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val 20 25 30

Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Leu Ser Leu Asp 35 40 45

Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln 50 60

Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His 65 70 75 80

Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr 85 90 95

Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys 100 105 110

Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp 115 120 125

Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe 130 140

Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile 145 150 155 160

Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr 165 170 175

04sequ.txt Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln 195 200 205 Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile 210 220 Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His 225 230 235 Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys 245 250 <21.0> 43 211> 1536 212> DNA Deinococcus radiodurans R1 <213> <220> <221> CDS <222> (1)..(1536)<223> atg ccg gat tac gac ctg atc gtc atg ggc gcg ggc cac aac gcg ctg Met Pro Asp Tyr Asp Leu Ile Val Met Gly Ala Gly His Asn Ala Leu 1 15 <400> 43 48 g act gct gcc tac gcc gcc cgg gcg ggc ctg aaa gtc ggc gtg ttc I Thr Ala Ala Tyr Ala Ala Arg Ala Gly Leu Lys Val Gly Val Phe 20 25 96 gag cgg cgg cac ctc gtc ggc ggg gcg gtc agc acc gag gag gtc gtg Glu Arg Arg His Leu Val Gly Gly Ala Val Ser Thr Glu Glu Val Val 35 40 45 144 ccc ggt tac cgc ttc gac tac ggc ggc agc gcc cac atc ctg att cgg Pro Gly Tyr Arg Phe Asp Tyr Gly Gly Ser Ala His Ile Leu Ile Arg 50 55 192

atg acg ccc atc gtg cgc gaa ctc gaa ctc acg cgg cac ggg ctg cat Met Thr Pro Ile Val Arg Glu Leu Glu Leu Thr Arg His Gly Leu His 65 70 75 80

tac ctc gaa gtg gac cct atg ttt cac gct tcc gac ggt gaa acg ccc Tyr Leu Glu Val Asp Pro Met Phe His Ala Ser Asp Gly Glu Thr Pro 85 90 95

tgg ttc att cac cgc gac gcc ggg cgg acc atc cgc gaa ctg gac gaa Trp Phe Ile His Arg Asp Ala Gly Arg Thr Ile Arg Glu Leu Asp Glu 100 110 Seite 54 240

288

336

a L	ag ys	ttt Phe	ccc Pro 115	ggg	g c / G	ag In	ggc Gly	gac Asp	gcc Ala 120	tac Tyr	g	gg (	cgc Arg	ttt Phe	Le	c (	gac Asp	gat Asp	: t	gg rp	384
a T	ca hr	ccc Pro 130	ttc Phe	gc Al	g c a A	gc .rg	gcc Ala	gtg Val 135	gcc Ala	gac Asp	C C	tg ' eu	ttc Phe	aac Asn 140	S	er .	gcg Ala	Pro	9 9	igg Sly	432
· P	cg ro .45	ctc Leu	gac Asp	ct Le	g g u G	igc Ty	aaa Lys 150	atg Met	gtg Val	ato Met	g c E A	rg	agc Ser 155	ggc	G	ag In	ggc Gly	aag Ly:	5 /	gac Asp 160	480
t	gg rp	aac Asn	gag	ca Gl	n L	etc Leu 165	ccg Pro	cgc Arg	atc Ile	cte Le	u A	gg krg L70	ccc Pro	tac Tyr	g G	gc ly	gac Asp	gt Va 17	1 /	gcg Ala	528
Å	gc Arg	gag Glu	tac Tyr	tt Ph 18	e s	agc Ser	gag Glu	gag Glu	cgc Arg	gt Va 18	<u> </u>	gg Arg	gct Ala	cco Pro	C C	tg eu	acc Thr 190	tg Tr	g p	atg Met	576
	gcg ala	gcc Ala	ca G1i 19	ı Se	ic (	ggc Gly	ccc Pro	cca Pro	ccc Pro 200	se	g g r A	gac Asp	ccg Pro	ct <u>e</u> Lei	ַ ג	gc er 205	gcg Ala	CC Pr	.c	ttt Phe	624
	ttg Leu	ctg Leu 210	Tr	g ca p Hi	is	ccg Pro	ctc Leu	tac Tyr 215	cac His	ga Gl	a (	ggc Gly	ggc Gly	gt Va 22	<u> </u>	icg \la	cgg Arg	CC Pr	°C	aaa Lys	672
	ggc Gly 225	Gly	ag Se	c gọ r G	jc ly	ggc Gly	ctg Leu 230	Thi	aaa Lys	a gc s Al	c a	ctg Leu	cgc Arg 235	, Ar	g g	gcc Ala	acc Thr	ga G	ag I u	gcc Ala 240	720
	gaa Glu	ggo	gg / Gl	c g y G	ag lu	gtc Val 245	Phe	ace Th	ga r Ası	g gc Al	g la	ccg Pro 250	gto Val	aa Lÿ	g (	gaa Glu	att Ile	L	tg eu 55	gtc Val	768
	aag Lys	gae S As	c gg o Gl	УĽ	ag ys 60	gcg Ala	cac Glr	g gg i Gl	c at	e Ai	9g 1g 65	ctg Leu	gaa Glu	a ag u Se	ic ir	ggc Gly	gag Gli 270		cg hr	tac Tyr	816
	aco Thi	gc Al	c cg a Ar 27	g A	cc la	gtc Val	gto Va	tc I Se	g gg r Gl 28	y Va	tc al	cac His	ate Il	c ct e Le	g eu	acc Thr 285	: In	t g r A	cg 1a	aat Asn	864
		c ct a Le 29	u Pr	c g	cc la	gaa Glu	ı ta ı Ty	t gt r Va 29	l Pr	t a	gc er	gcc Ala	gc Al	a Al	99 79 00	aat Asr	gt Va	g c l A	gc	gtg Val	912
	gg G1: 30	y As	c gọ n G	ic t	tc he	ggo Gly	at Me 31	t II	t tt e Le	g c u A	gc rg	cto	gc Al 31	a Lo	t <b>c</b> eu	agt Sei	t ga r Gl	a a u L	aa .ys	gtc Val 320	960
	aa Ly	a ta s Ty	c c	gt d rg H	ac Iis	cae His 32	<u>s</u> Th	c ga r Gl	g co u Pr	c g	ac sp	tca Ser 330	Ar	c a g I	tc le	gg Gl	c ct y Le	u	gga Sly 335	ttg Leu	1008
	ct Le	g at u Il	c a e L	ys /	aac Asn 340	Gl	g cg u Ar	g ca g G	a a <u>ı</u> n I	le №	itg let 845	Gli	g gg n Gl	c t y T	ac yr	gg G l	c ga y G1 35	u	tac Fyr	ctc Leu	1056
	gc Al	c gg a G	y G	ag ( 1n   55	ccc Pro	ac Th	c ac r Th	c ga	sp Pi	cg c ro F 60	cc Pro	ct Le	c gt u Va	c g	cc la	at Me 36	τ 56	jc 1 er 1	tto Pho	c agc e Ser	1104
	gc Al	a Va	g g al A 70	ac sp	gac Asp	tc Se	g ct r Le	eu A	cc co la P 75	ca d ro F	cg Pro	aa As	c gg n G	ıy A	ac sp 80	٧a	g ti	g :	tg: Tr	g ctg p Leu	1152
												Se	ite	55							

								ctc Leu							acg Thr 400	1200
cgc Arg	acc Thr	gcc Ala	gaa Glu	gcg Ala 405	cgg Arg	gag Glu	aac Asn	atc Ile	ctg Leu 410	cgg Arg	gcc Ala	ttt Phe	gag Glu	cac His 415	tac Tyr	1248
gcg Ala	ccg Pro	ggc Gly	acc Thr 420	cgc Arg	gac Asp	acg Thr	att Ile	gtg Val 425	ggc Gly	gaa Glu	ctc Leu	gtg Val	cag Gln 430	acg Thr	ccg Pro	1296
cag Gln	tgg Trp	ctg Leu 435	gaa Glu	acc Thr	aac Asn	ctc Leu	ggc Gly 440	ctg Leu	cac His	cgg Arg	ggc Gly	aac Asn 445	gtg Val	atg Met	cac His	1344
ctg Leu	gaa Glu 450	atg Met	tcc Ser	ttc Phe	gac Asp	cag Gln 455	atg Met	ttc Phe	tcc Ser	ttc Phe	cgc Arg 460	ccc Pro	tgg Trp	ctg Leu	aaa Lys	1392
gcg Ala 65	agc Ser	cag Gln	tac Tyr	cgc Arg	tgg Trp 470	ccg Pro	ggc Gly	gtg Val	cag Gln	ggg Gly 475	ctg Leu	tac Tyr	ctc Leu	acc Thr	ggc Gly 480	1440
gcc Ala	agc Ser	acc Thr	cac His	ccc Pro 485	ggc Gly	gga Gly	ggc Gly	atc Ile	atg Met 490	ggc Gly	gcc Ala	tcg Ser	gga Gly	cgc Arg 495	aac Asn	1488
	gcg Ala			atc Ile	gtg Val	aag Lys	gac Asp	ctg Leu 505	acg Thr	cgg Arg	agg Arg	cgc Arg	tgg Trp 510	aaa Lys	tga	1536

<210> 44

<211> 511

<212> PRT

<213> Deinococcus radiodurans R1

<400> 44

t Pro Asp Tyr Asp Leu Ile Val Met Gly Ala Gly His Asn Ala Leu 5 10 15

Val Thr Ala Ala Tyr Ala Ala Arg Ala Gly Leu Lys Val Gly Val Phe 20 30

Glu Arg Arg His Leu Val Gly Gly Ala Val Ser Thr Glu Glu Val Val 35 40 45

Pro Gly Tyr Arg Phe Asp Tyr Gly Gly Ser Ala His Ile Leu Ile Arg 50 60

Met Thr Pro Ile Val Arg Glu Leu Glu Leu Thr Arg His Gly Leu His 65 70 75 80

Tyr Leu Glu Val Asp Pro Met Phe His Ala Ser Asp Gly Glu Thr Pro 85 90 95
Seite 56

Trp Phe Ile His Arg Asp Ala Gly Arg Thr Ile Arg Glu Leu Asp Glu 100 105 110 Lys Phe Pro Gly Gln Gly Asp Ala Tyr Gly Arg Phe Leu Asp Asp Trp 115 120 125 Thr Pro Phe Ala Arg Ala Val Ala Asp Leu Phe Asn Ser Ala Pro Gly 130 140 Pro Leu Asp Leu Gly Lys Met Val Met Arg Ser Gly Gln Gly Lys Asp 145 150 155 160 Trp Asn Glu Gln Leu Pro Arg Ile Leu Arg Pro Tyr Gly Asp Val Ala 165 170 175 Arg Glu Tyr Phe Ser Glu Glu Arg Val Arg Ala Pro Leu Thr Trp Met 180 185 190 Ala Ala Gln Ser Gly Pro Pro Pro Ser Asp Pro Leu Ser Ala Pro Phe 195 200 205 Leu Leu Trp His Pro Leu Tyr His Glu Gly Gly Val Ala Arg Pro Lys 210 215 220 Gly Gly Ser Gly Gly Leu Thr Lys Ala Leu Arg Arg Ala Thr Glu Ala 225 230 240 Glu Gly Gly Glu Val Phe Thr Asp Ala Pro Val Lys Glu Ile Leu Val 245 250 255 Lys Asp Gly Lys Ala Gln Gly Ile Arg Leu Glu Ser Gly Glu Thr Tyr 260 265 270 r Ala Arg Ala Val Val Ser Gly Val His Ile Leu Thr Thr Ala Asn 275 280 285 Ala Leu Pro Ala Glu Tyr Val Pro Ser Ala Ala Arg Asn Val Arg Val 290 295 300 Gly Asn Gly Phe Gly Met Ile Leu Arg Leu Ala Leu Ser Glu Lys Val 305 310 . 315 320 Lys Tyr Arg His His Thr Glu Pro Asp Ser Arg Ile Gly Leu Gly Leu 325 330 335 Leu Ile Lys Asn Glu Arg Gln Ile Met Gln Gly Tyr Gly Glu Tyr Leu 340 345 350 Ala Gly Gln Pro Thr Thr Asp Pro Pro Leu Val Ala Met Ser Phe Ser 355 360 365

Seite 57

Ala	Va I 370	Asp	Asp	Ser	Leu	375	Pro	Pro	ASN	Gly	Asp 380	Val	Leu	Trp	Leu		
Trp 385	Ala	Gln	Tyr	Tyr	Pro 390	Phe	Glu	Leu	Ala	Thr 395	Gly	ser	Trp	Glu	Thr 400		
Arg	Thr	Аlа	Glu	Ala 405	Arg	Glu	Asn	Ile	Leu 410	Arg	Ala	Phe	Glu	ніs 415	туг		
Ala	Pro	Gly	Thr 420	Arg	Asp	Thr	Ile	Va7 425	GÌу	Glu	Leu	Val	G]n 430	Thr	Pro		
Gln	Trp	Leu 435	Glu	Thr	Asn	Leu	Gly 440	Leu	His	Arg	Gly	Asn 445	٧a٦	Met	His		
eu	G]u 450	Met	Ser	Phe	Asp	G]n 455	Met	Phe	Ser	Phe	Arg 460	Pro	Тгр	Leu	Lys		
Ala 465	Ser	Gln	Tyr	Arg	Trp 470	Pro	Gly	Val	Gln	Gly 475	Leu	Tyr	Leu	Thr	Gly 480		
Ala	Ser	Thr	His	Pro 485	Gly	Gly	Gly	Ile	Met 490		Ala	Ser	Gly	Arg 495	Asn		
Ala	Αla	Arg	Va1 500	Ile	٧a٦	Lys	Asp	Leu 505	Thr	Arg	Arg	Arg	Trp 510				
<21	0>	45															
<21	1>	1650															
<21	2>	DNA															
<21		Lyco	pers	icon	esc	ulen	tum							,			
<22 <22	0> 1>	CDS															
<22			) (	1614	`												
<22		(112	) · · (	1017	,												
<40		45															
_	-		aact	tttc	tc t	cttc	acta	g ct	gttt	acat	gct	tgaa	att	tcaa	gatttt	•	60
agg	accc	cat	ttga	agtt	tt c	ttga	aaca	a at	atta	.ccct	gtt	:ggaa	aaa		g gat t Asp	1.7	17
act Thr	ttg Leu	ttg Leu 5	aaa Lys	acc Thr	cca Pro	aat Asn	aac Asn 10	ctt Leu	Glu	ttt Phe	Leu	aad Asr 15	cca Pro	cat His	cat His	10	65
									ンセリ	LC .	, 0						

ggt Gly	ttt Phe 20	gct Ala	gtt Val	aaa Lys	gct Ala	agt ser 25	acc Thr	ttt Phe	aga Arg	tct Ser	gag Glu 30	aag Lys	cat His	cat His	aat Asn	213
ttt Phe 35	Gly	tct Ser	agg Arg	aag Lys	ttt Phe 40	tgt Cys	gaa Glu	act Thr	ttg Leu	ggt Gly 45	aga Arg	agt Ser	gtt Val	tgt Cys	gtt Val 50	261
aag Lys	ggt Gly	agt Ser	agt Ser	agt Ser 55	gct Ala	ctt Leu	tta Leu	gag Glu	ctt Leu 60	gta Val	cct Pro	gag Glu	acc Thr	aaa Lys 65	aag Lys	309
					gag Glu											357
gtt Val	gtg Val	gat Asp 85	ctt Leu	gct Ala	gtg Val	gtt Val	ggt Gly 90	ggt Gly	ggc Gly	cct Pro	gca Ala	gga Gly 95	ctt Leu	gct Ala	gtt Val	405
gca Ala	cag Gln 100	caa Gln	gtt Val	tct Ser	gaa Glu	gca Ala 105	gga Gly	ctc Leu	tct Ser	gtt val	tgt Cys 110	tca Ser	att Ile	gat Asp	ccg Pro	453
aat Asn 115	cct Pro	aaa Lys	ttg Leu	ata Ile	tgg Trp 120	cct Pro	aat Asn	aac Asn	tat Tyr	ggt Gly 125	gtt Val	tgg Trp	gtg Val	gat Asp	gaa Glu 130	501
ttt Phe	gag Glu	gct Ala	atg Met	gac Asp 135	ttg Leu	tta Leu	gat Asp	tgt Cys	cta Leu 140	gat Asp	gct Ala	acc Thr	tgg Trp	tct Ser 145	ggt Gly	549
gca Ala	gca Ala	gtg Val	tac Tyr 150	att Ile	gat Asp	gat Asp	aat Asn	acg Thr 155	gct Ala	aaa Lys	gat Asp	ctt Leu	cat His 160	aga Arg	cct Pro	597
					cgg Arg			Leu							aaa Lys	645
		Met			gtt Val							٧a٦				693
e 5	cat His	gag Glu	gaa Glu	tcg Ser	aaa Lys 200	Ser	atg Met	ttg Leu	ata Ile	tgc Cys 205	Asn	gat Asp	ggt Gly	att Ile	act Thr 210	741
att Ile	cag Gln	gca Ala	acg Thr	gtg Val 215	Val	ctc Leu	gat Asp	gca Ala	act Thr 220	Gly	ttc Phe	tct Ser	aga Arg	tct Ser 225	ctt Leu	789
gtt Val	cag Gln	tat Tyr	gat Asp 230	Lys	cct Pro	tat Tyr	aac Asn	ccc Pro 235	Gly	tat Tyr	caa Glr	gtt Val	gct Ala 240	Tyr	ggc	837
att Ile	ttg Leu	gct Ala 245	Glu	gtg Val	gaa Glu	gag Glu	cac His 250	Pro	ttt Phe	gat Asp	gta Val	aac Asn 255	Lys	at <u>c</u> Met	gtt Val	885
ttc Phe	atg Met 260	Asp	tgg Trp	cga Arg	gat Asp	tct Ser 265	His	ttg Leu	aag Lys	aac Asn	aat Asr 270	Thr	gat Asp	cto Leu	aag Lys	933
gag G1u 275	Arg	aat Asr	agt Ser	aga Arg	ata 11e 280	Pro	act Thr	ttt Phe	. Leu	tat Tyr 285	Āla	a atg a Met	cca Pro	ttt Phe	tca Ser 290	981

# O4sequ.txt

tcc Ser	aac Asn	agg Arg	ata Ile	ttt Phe 295	ctt Leu	gaa Glu	gaa Glu	aca Thr	tca Ser 300	ctc Leu	gta Val	gct Ala	cgt Arg	cct Pro 305	ggc Gly		1029
ttg Leu	cgt Arg	ata Ile	gat Asp 310	gat Asp	att Ile	caa Gln	gaa Glu	cga Arg 315	atg Met	gtg Val	gct Ala	cgt Arg	tta Leu 320	aac Asn	cat His		1077
ttg Leu	ggg Gly	ata 11e 325	aaa Lys	gtg Val	aag Lys	agc Ser	att Ile 330	gaa Glu	gaa Glu	gat Asp	gaa Glu	cat His 335	tgt Cys	cta Leu	ata Ile		1125 ·
cca Pro	atg Met 340	ggt Gly	ggt Gly	cca Pro	ctt Leu	cca Pro 345	gta Val	tta Leu	cct Pro	cag Gln	aga Arg 350	gtc Val	gtt Val	gga Gly	atc Ile		1173
ggt Gly 355	ggt Gly	aca Thr	gct Ala	ggc Gly	atg Met 360	gtt Val	cat His	cca Pro	tcc Ser	acc Thr 365	ggt Gly	tat Tyr	atg Met	gtg Val	gca Ala 370		1221
agg rg	aca Thr	cta Leu	gct Ala	gcg Ala 375	gct Ala	cct Pro	gtt Val	gtt Val	gcc Ala 380	aat Asn	gcc Ala	ata Ile	att Ile	caa Gln 385	tac Tyr		1269
ctc Leu	ggt Gly	tct Ser	gaa Glu 390	aga Arg	agt Ser	cat His	tcg Ser	ggt Gly 395	aat Asn	gaa Glu	tta Leu	tcc Ser	aca Thr 400	gct Ala	gtt Val		1317
						ata Ile											1365
		ĞĪy				ctt Leu 425									aga Arg	•	1413
agg Arg 435	ttc Phe	ttt Phe	gat Asp	gca Ala	ttc Phe 440	Phe	gac Asp	tta Leu	gaa Glu	cct Pro 445	Arg	tat Tyr	tgg Trp	cat His	ggc Gly 450		1461
ttc Phe	tta Leu	tcg Ser	tct Ser	cga Arg 455	ttg Leu	ttt Phe	cta Leu	cct Pro	gaa Glu 460	Leu	ata Ile	gtt Val	Phe	ggg Gly 465	ctg Leu		1509
t				His					Ser					Met	aca Thr		1557
aag Lys	gga Gly	act Thr 485	٧a٦	cca Pro	tta Leu	gta Val	aat Asn 490	Met	atc Ile	aac Asn	aat Asn	ttg Leu 495	Leu	cag Gln	gat Asp		1605
	gaa Glu 500	ı -	atc	cgag	taa	ttcg	gaat	ct t	gtcc	aato	t cg	tgco	:				1650
<21	.0>	46		•													
<21	.1>	500															
.21	٦.																

<212> PRT

<213> Lycopersicon esculentum

<400> 46

Met Asp Thr Leu Leu Lys Thr Pro Asn Asn Leu Glu Phe Leu Asn Pro 15 Pro His His Gly Phe Ala Val Lys Ala Ser Thr Phe Arg Ser Glu Lys His 30 Lys His Asn Phe Gly Ser Arg Lys Phe Cys Glu Thr Leu Gly Arg Ser Val Cys Val Lys Gly Ser Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys 70 Sely Val Val Val Asp Leu Ala Val Val Gly Gly Gly Pro Ala Gly Leu 95 Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys 80 Sely Val Val Val Asp Leu Ala Val Val Gly Gly Gly Pro Ala Gly Leu 95 Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys 80 Sely Val Val Val Val Asp Leu Ala Val Val Gly Gly Gly Pro Ala Gly Leu 95 Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys 80 Sely Val Val Val Val Asp Leu Ala Val Val Gly Gly Gly Pro Ala Gly Leu 95 Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys 80 Sely Val Val Val Val Asp Leu Ala Val Val Gly Gly Gly Pro Ala Gly Leu

Ala Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile 100 105 110

Asp Pro Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val 115 120 125

Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp 130 135 140

Ser Gly Ala Ala Val Tyr Ile Asp Asp Asn Thr Ala Lys Asp Leu His 145 150 150 160

Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Met 165 170 175

n Lys Cys Ile Met Asn Gly Val Lys Phe His Gln Ala Lys Val Ile 180 185 190

Lys Val Ile His Glu Glu Ser Lys Ser Met Leu Ile Cys Asn Asp Gly 195 200 205

Ile Thr Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly Phe Ser Arg 210 215 220

Ser Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala 225 230 235 240

Tyr Gly Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp Val Asn Lys 245 250 255

Met Val Phe Met Asp Trp Arg Asp Ser His Leu Lys Asn Asn Thr Asp 260 265 270 Seite 61

Leu Lys Glu Arg Asn Ser Arg Ile Pro Thr Phe Leu Tyr Ala Met Pro 275 280 285

Phe Ser Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg 290 295 300

Pro Gly Leu Arg Ile Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu 305 310 315

Asn His Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys 325 330 335

Leu Ile Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val 340 345 350

Gly Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met 355 360 365

Val Ala Arg Thr Leu Ala Ala Ala Pro Val Val Ala Asn Ala Ile Ile 370 375 380

Gln Tyr Leu Gly Ser Glu Arg Ser His Ser Gly Asn Glu Leu Ser Thr 385 390 395 400

Ala Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Gln Arg Glu 410 415

Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Pro Ala 420 425 430

Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp 435 440 445

Gly Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Ile Val Phe 450 460

Gly Leu Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Phe Glu Ile 465 470 475 480

Met Thr Lys Gly Thr Val Pro Leu Val Asn Met Ile Asn Asn Leu Leu 485 490 495

Gln Asp Lys Glu 500

<210> 47

<211> 1666

<212> DNA

<213> Lycopersicon esculentum

<220>

<221> CDS

<222> (1)..(1494)

<223>

	atα	> 4 gaa Glu	act	ctt Leu	ctc Leu 5	aag Lys	cct Pro	ttt Phe	cca Pro	tct Ser 10	ctt Leu	tta Leu	ctt Leu	tcc Ser	tct Ser 15	cct Pro		48
	aca Thr	ccc Pro	cat His	agg Arg 20	tct Ser	att Ile	ttc Phe	caa Gln	caa Gln 25	aat Asn	ccc Pro	tct Ser	ttt Phe	cta Leu 30	agt Ser	ccc Pro		96
	cc hr	acc Thr	aaa Lys 35	aaa Lys	aaa Lys	tca Ser	aga Arg	aaa Lys 40	tgt Cys	ctt Leu	ctt Leu	aga Arg	aac Asn 45	aaa Lys	agt Ser	agt Ser		144
	aaa Lys	ctt Leu 50	ttt Phe	tgt Cys	agc Ser	ttt Phe	ctt Leu 55	gat Asp	tta Leu	gca Ala	ccc Pro	aca Thr 60	tca Ser	aag Lys	cca Pro	gag Glu		192
	tct Ser 65	tta Leu	gat Asp	gtt Val	aac Asn	atc Ile 70	tca Ser	tgg Trp	gtt Val	gat Asp	cct Pro 75	aat Asn	tcg Ser	aat Asn	cgg Arg	gct Ala 80		240
	caa Gln	ttc Phe	gac Asp	gtg Val	atc Ile 85	att Ile	atc Ile	gga Gly	gct Ala	ggc Gly 90	cct Pro	gct Ala	ggg Gly	ctc Leu	agg Arg 95	cta Leu		288
	gct Ala	gaa Glu	caa Gln	gtt Val 100	Ser	aaa Lys	tat Tyr	ggt Gly	att Ile 105	aag Lys	gta Val	tgt Cys	tgt Cys	gtt Val 110	gac Asp	cct Pro		336
4	tca Ser	cca Pro	ctc Leu 115	Ser	atg Met	tgg Trp	cca Pro	aat Asn 120	Asn	tat Tyr	ggt Gly	gtt Val	tgg Trp 125	gtt Val	gat Asp	gag Glu	•	384
	Phe	gag Glu 130	Asn	tta Leu	gga Gly	ctg Leu	gaa Glu 135	Asn	tgt Cys	tta Leu	gat Asp	cat His 140	Lys	tgg Trp	cct Pro	atg Met		432
	act Thr 145	· Cys	gtg Val	cat His	ata Ile	aat Asn 150	Asp	aac Asn	aaa Lys	act Thr	aag Lys 155	Tyr	ttg Leu	gga Gly	aga ⁄Arg	cca Pro 160		480
	tat Tyr	ggt Gly	aga / Arg	gtt Val	agt Ser 165	· Arg	aag Lys	aag Lys	ctg Leu	aag Lys 170	Leu	aaa Lys	ttg Leu	ttg Lei	aat Asn 175	agt Ser		528
	tgt Cys	gtt Va	gag I Gli	aac Asr 180	1 Arg	gtg Val	l aag Lys	ttt Phe	tat Tyr 185	` Lys	gct Ala	aag Lys	gtt Val	tgg Tri 190	o Lys	gtg Val		576
	gaa Glu	a cat u His	t gaa s Glu 195	ı Glι	ttt Phe	gag Glu	tct Ser	tca Ser 200	. I16	gtt val	tgt Cys	gat S Asp	gat Asp 205	) GI	t aag y Lys	aag Lys		624
	ata	a aga	a gg1	t ag1	t ttg	g gtt	gtg	g gat	gca		gg1		c gc1	t ag	t gai	ttt		672

ıle	Arg 210	GΊу	Ser	Leu		Val 215	Asp		4seq Ser	Gly		Ala	ser .	Asp	Phe	
ata Ile 225	gag Glu	tat Tyr	gac Asp	Arg	cca Pro 230	aga Arg	aac Asn	cat His	Gly	tat Tyr 235	caa Gln	att Ile	gct Ala	cat His	ggg Gly 240	720
gtt Val	tta Leu	gta Val	gaa Glu	gtt Val 245	gat Asp	aat Asn	cat His	cca Pro	ttt Phe 250	gat Asp	ttg Leu	gat Asp	aaa Lys	atg Met 255	gtg Val	768
ctt Leu	atg Met	gat Asp	tgg Trp 260	agg Arg	gat Asp	tct Ser	cat His	ttg Leu 265	ggt Gly	aat Asn	gag Glu	cca Pro	tat Tyr 270	tta Leu	agg Arg	816
gtg Val	aat Asn	aat Asn 275	gct Ala	aaa Lys	gaa Glu	cca Pro	aca Thr 280	ttc Phe	ttg Leu	tat Tyr	gca Ala	atg Met 285	cca Pro	ttt Phe	gat Asp	864
aga Arg	gat Asp 290	ttg Leu	gtt Val	ttc Phe	ttg Leu	gaa Glu 295	gag Glu	act Thr	tct Ser	ttg Leu	gtg Val 300	agt Ser	cgt Arg	cct Pro	gtt Val	912
ta 2u 305	tcg Ser	tat Tyr	atg Met	gaa Glu	gta Val 310	aaa Lys	aga Arg	agg Arg	atg Met	gtg Val 315	gca Ala	aga Arg	tta Leu	agg Arg	cat His 320	960
ttg Leu	ggg Gly	atc Ile	aaa Lys	gtg Val 325	aaa Lys	agt Ser	gtt Val	att Ile	gag Glu 330	gaa Glu	gag Glu	aaa Lys	tgt Cys	gtg Val 335	atc Ile	1008
cct Pro	atg Met	gga Gly	gga Gly 340	Pro	ctt Leu	ccg Pro	cgg Arg	att Ile 345	cct Pro	caa Gln	aat Asn	gtt Val	atg Met 350	gct Ala	att Ile	1056
ggt Gly	ggg Gly	aat Asn 355	Ser	ggg Gly	ata Ile	gtt Val	cat His 360	Pro	tca Ser	aca Thr	ggg GTy	tac Tyr 365	Met	gtg Val	gct Ala	1104
agg Arg	agc Ser 370	Met	gct Ala	tta Leu	gca Ala	cca Pro 375	Val	cta Leu	gct Ala	gaa Glu	gcc Ala 380	atc lle	gtc Val	gag Glu	Gly	1152
ctt Leu	เดิโร	tca Ser	a aca Thr	aga Arg	atg Met 390	: Ile	aga Arg	ggg Gly	tct Ser	caa Gln 395	Lei	tac ı Tyr	cat His	aga Arg	gtt Val 400	1200
ır	aat Asr	ggt Gly	t ttg / Lei	tgg Trp 405	Pro	ttg Leu	gat I Asp	aga Arg	aga Arg 410	Cys	gti Va	t aga l Arg	gaa Glu	tgt Cys 41	tat Tyr	1248
tca Sei	a ttt Phe	ggg Gly	g ato y Mei 420	G G T L	g aca u Thr	ttg Lei	j ttg i Lei	g aag 1 Lys 425	Leu	gat Asp	tte Lei	g aaa u Lys	ggg Gly 430	/ Thi	agg Arg	1296
aga Arg	a ttg g Lei	tti Pho 43	e Āsi	c gct o Ala	t tto a Pho	tt1 e Phe	ga 2 As <sub>1</sub> 440	<u>Le</u>	t gat u Asp	cc1 Pro	t aaa o Ly:	a tad s Tyr 449	Tr	g caa Gli	aggg Gly	1344
tte Pho	c cti e Lei 450	ı se	t tc r Se	a aga r Ar	a ttg g Lei	g tci u Sei 45	r Va	c aaa l Ly:	a gaa s Glu	ı cti ı Lei	t gg u G1 46	y Lei	a cto	c ag u Se	c ttg r Leu	1392
tg Cy: 46	s Lei	t tt u Ph	c gg e Gl	a ca y Hi:	t gg s Gly 47	y se	a aa r As	c at n Me	g act t Thi	t agg	g Le	g ga¹ u As <sub>l</sub>	t at	t gt e Va	t aca 1 Thr 480	
aa	a tg	t cc	t ct	t cc	t tt	g gt	t ag	a ct		t gg ite		t ct	a gc	a at	a gag	1488

O4sequ.txt

Lys Cys Pro Leu Pro Leu Val Arg Leu Ile Gly Asn Leu Ala Ile Glu
485 agc ctt tgaatgtgaa aagtttgaat cattttcttc attttaattt ctttgattat 1544 ser Leu 1604 tttcatattt tctcaattgc aaaagtgaga taagagctac atactgtcaa caaataaact 1664 actattggaa agttaaaata tgtgtttgtt gtatgttatt ctaatggaat ggattttgta 1666 <210> 48 498 <211> <212> PRT <213> Lycopersicon esculentum 400> 48 Met Glu Ala Leu Leu Lys Pro Phe Pro Ser Leu Leu Leu Ser Ser Pro 1 10 15 Thr Pro His Arg Ser Ile Phe Gln Gln Asn Pro Ser Phe Leu Ser Pro 20 25 30 Thr Thr Lys Lys Ser Arg Lys Cys Leu Leu Arg Asn Lys Ser Ser 35 40 45Lys Leu Phe Cys Ser Phe Leu Asp Leu Ala Pro Thr Ser Lys Pro Glu 50 60 Ser Leu Asp Val Asn Ile Ser Trp Val Asp Pro Asn Ser Asn Arg Ala 65 70 75 80 h Phe Asp Val Ile Ile Ile Gly Ala Gly Pro Ala Gly Leu Arg Leu 85 90 95 Ala Glu Gln Val Ser Lys Tyr Gly Ile Lys Val Cys Cys Val Asp Pro 100 105 110

Ser Pro Leu Ser Met Trp Pro Asn Asn Tyr Gly Val Trp Val Asp Glu
Phe Glu Asn Leu Gly Leu Glu Asn Cys Leu Asp His Lys Trp Pro Met
Thr Cys Val His Ile Asn Asp Asn Lys Thr Lys Tyr Leu Gly Arg Pro
145

Tyr Gly Arg Val Ser Arg Lys Leu Lys Leu Lys Leu Leu Asn Ser 165 170 175 Seite 65

Cys Val Glu Asn Arg Val Lys Phe Tyr Lys Ala Lys Val Trp Lys Val 180 185 190 Glu His Glu Glu Phe Glu Ser Ser Ile Val Cys Asp Gly Lys Lys 195 200 205 Ile Arg Gly Ser Leu Val Val Asp Ala Ser Gly Phe Ala Ser Asp Phe 210 220 Ile Glu Tyr Asp Arg Pro Arg Asn His Gly Tyr Gln Ile Ala His Gly 225 230 235 240 Val Leu Val Glu Val Asp Asn His Pro Phe Asp Leu Asp Lys Met Val 245 250 255 eu Met Asp Trp Arg Asp Ser His Leu Gly Asn Glu Pro Tyr Leu Arg 260 265 270 Val Asn Asn Ala Lys Glu Pro Thr Phe Leu Tyr Ala Met Pro Phe Asp 275 280 285 Arg Asp Leu Val Phe Leu Glu Glu Thr Ser Leu Val Ser Arg Pro Val 290 295 300 Leu Ser Tyr Met Glu Val Lys Arg Arg Met Val Ala Arg Leu Arg His 305 310 315 320 Leu Gly Ile Lys Val Lys Ser Val Ile Glu Glu Glu Lys Cys Val Ile 325 330 335 Pro Met Gly Gly Pro Leu Pro Arg Ile Pro Gln Asn Val Met Ala Ile 340 345 350 y Gly Asn Ser Gly Ile Val His Pro Ser Thr Gly Tyr Met Val Ala 355 360 365 Arg Ser Met Ala Leu Ala Pro Val Leu Ala Glu Ala Ile Val Glu Gly 370 375 380 Leu Gly Ser Thr Arg Met Ile Arg Gly Ser Gln Leu Tyr His Arg Val 385 390 395 400 Trp Asn Gly Leu Trp Pro Leu Asp Arg Arg Cys Val Arg Glu Cys Tyr 405 410 415 Ser Phe Gly Met Glu Thr Leu Leu Lys Leu Asp Leu Lys Gly Thr Arg 420 425 430 Arg Leu Phe Asp Ala Phe Phe Asp Leu Asp Pro Lys Tyr Trp Gln Gly 435 445 Seite 66

٢	he	Leu 450	se	r S	er A	rg L	_eu	ser 455	val	Lys	Glu	Leu	Gly 460	Leu	ı Le	eu S	Ser	Leu		
4	ys 165	Leu	Ph	e G	Ίу∙н	is (	31y 470	Ser	Asn	Met	Thr	Arg 475	Leu	Asp	o I	le v	/al	Thr 480		
ŧ	.ys	Cys	Pr	o L	eu F	ro 1 185	Leu	۷al	Arg	Leu	11e 490	Gly	Asn	Lei	u A	la :	11e 495	Glu		
9	ser	Leu	I																	
	<210	0>	49																	
	<21	1>	160	8																
	<b>&lt;21</b>	2>	DNA																	
	21	3>	нае	emat	toco	ccus	pl	uvia	lis										-	
																			•	
	<22	0>																		
	<22	1>	CD:	5																
	<22	2>	(3)	)	(971	)														
		٦.																		
	<22	.3>							•											
	<22	.3>							·											
	<40	00> aca Thr	49 tt Ph	t c e H	ac a is L	ag (	ccc Pro	gtg Val	agc Ser	ggt Gly	gca Ala	agc Ser 10	gct Ala	ct; Lei	g C	cc (	cac His	atc Ile 15	47	
	<40 ct	00> aca Thr 1	tt Ph	e H ct	is L	ys ! cat	Pro 5 cto	Val cat	Ser ca	Gly a tc	Ala a tt	Ser	Ala t qc	Lei t a	u P cc	ro i acg	atg	15 ct	g . 95	
	<40 ct ggo Gly	00> aca Thr 1 cc/Pr	tt Ph a c o P	e H ct ro	is L cct Pro	cat His 20	cto Leu	Val cat His	Ser c cgg s Arg	Gly g tc g se	Ala a tt r Ph 25 g gc s Al	Ser 10 t ac	Ala t gc a Al	Lei ta a Ti c g	cc hr tt al	acg Thr	ato Mei 30	15 g ctg t Le	g 95 u 143	
	<40 ct ggc Gly ser	00> aca Thr 1 c cc / Pr	a co P g co	e H ct ro tg eu	cct Pro cag Gln 35	cat His 20 tca Ser	cto Leu ato	Val : cat : Hi: : age : Se	Ser c gt c gt	Gly g tc g se c aa l Ly 40	a ttr Ph 25 g gc s Al	Ser 10 t gc e Ala	t gc a Al c cg g Ar	t a T C g V	cc hr tt al	acg Thr gaa Glu 45	atg Mei 30 cta Lei	11e 15 c te a gc	g 95 u 143 a 191	
	<400 ct	00> aca Thr 1 ccc/Pr aar Ly	Ph a C P C I C A C P I C P I C A C P I C P I C A C P I C P I C A C P I C A C P I C A C P I C A C P I C A C P I C A C P I C A C P I C A C P I C A C P I C A C P I C A C P I C A C P I C A C	ct ro tg eu	cct Pro cag Gln 35 acg Thr	cat His 20 tca Ser . cgg Arg	ctc Leu atc Ile ccc	val cat His age Se Ly	ser c gt c gt c gt c ya a gt s 55 a gc a Al	g tc. g se c aa l Ly 40 c tg l Cy	a ttrph 25 g gc s Al	t gc e Ala c cg a Ar	t gc a Al c cg g Ar t gc	tad a TI c g V ta G G	cc hr tt al ag	acg Thr gaa Glu 45 cgg Arg	ate Med 30 cta Lei tg	g cter g cter a gc u Al	g 95 c 143 g 191 r 239	
	cgc Arg	00> aca Thr 1 cc Pr aay Ly c ga su Va 65 c gt	Ph a CP g C L c a c p S c c c c c c c c c c c c c c c c c	e H ct ro tg eu tc le cle cgg	cct Pro cag Gln 35 acg Thr ctg Leu	cat His 20 tca Ser . cgg Arg	ctc Leu atcc Ile ccc Pro	cat His age Ses Ly g gc l Al 70	ser c gg c gt r Va a gt s Va a gc a Al	g tc. g se c aaal Ly 40 c tg a cca	a ttr r Ph 25 g gc s Al c ct ca cal	t gc e Ala c cg a Ar g ca u Hi	t gc a Al c cg g Ar t gc s a ga r 75	taa cgy ca 6 gg c	cchr ttal agln agu	acg Thr gaa GJu 45 cgg Arg gcg Ala	atcomes at the second s	g ct. s ct.	g 95 c 143 a 191 r 239 y 287	
	<400 ct ggGGl) cgGArd ttta	00> aca Thr 1 cc Pr aau ya Gg As cc gt Cc Cc	Ph a C P C I C A C I C I C I C I C I C I C I C I	e H ct ro tg eu tce le gg rg	cct Pro cag Gln 35 acg Thr ctg Leu gct	cat His 20 tca Ser .cgg Arg cga Arg	ctc Leu atc ccc Pro gt, Va gg; SGI: 85	cat His age Se Se Caa D Ly G GC G Al	ser t cgg t Arg t ya a gta a ya a gg a Al gg a Gl t at	g tc. g se c aa l Ly c tg a cc a Pr c ga	a tthr 25 g gc s Al c cts ca Gl	t gc: e Al: c cg a Ar: g ca u Hi g ca u Hi g ca u Hi	t gCi t a Al c cg Ar t a Gl a c Se	ta cg ta gu gar d	cchr ttal agno agu	acg Thr gaaa Glu 45 cgg Ala gat cgg	atc Mei 30 cta Lei tg Cy	g ct. gc Le gc Le gg Gl a gc Ly	g 95 c 143 a 191 r 239 y 287 a 335	•

04sequ.txt Arg Glu Gln Leu Ser Tyr Gln Ala Ala Ala Ile Ala Ala Ser Ile Gly gtg tca ggc att gcc atc ttc gcc acc tac ctg aga ttt gcc atg cac Val Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His 431 atg acc gtg ggc gca gtg cca tgg ggt gaa gtg gct ggc act ctc Met Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu 145 150 155 479 ctc ttg gtg gtt ggt ggc gcg ctc ggc atg gag atg tat gcc cgc tat Leu Leu Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr 160 165 170 175 527 gca cac aaa gcc atc tgg cat gag tcg cct ctg ggc tgg ctg ctg cac Ala His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His 180 185 575 aag agc cac cac aca cct cgc act gga ccc ttt gaa gcc aac gac ttg Lys Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu 195 200 205 623 rt gca atc atc aat gga ctg ccc gcc atg ctc ctg tgt acc ttt ggc he Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly 210 220 671 719 ttc tgg ctg ccc aac gtc ctg ggg gcg gcc tgc ttt gga gcg ggg ctg Phe Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu 225 230 235 ggc atc acg cta tac ggc atg gca tat atg ttt gta cac gat ggc ctg Gly Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu 240 255 250 767 gtg cac agg cgc ttt ccc acc ggg ccc atc gct ggc ctg ccc tac atg Val His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met 260 265 270 815 aag cgc ctg aca gtg gcc cac cag cta cac cac agc ggc aag tac ggt Lys Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly 275 280 285 863 911 ggc gcg ccc tgg ggt atg ttc ttg ggt cca cag gag ctg cag cac att Gly Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile 290 295 300 Ca ggt gcg gcg gag gag gtg gag cga ctg gtc ctg gaa ctg gac tgg Pro Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp 305 310 315 959 1011 tcc aag cgg tag ggtgcggaac caggcacgct ggtttcacac ctcatgcctg Ser Lys Arg 320 1071 tgataaggtg tggctagagc gatgcgtgtg agacgggtat gtcacggtcg actggtctga 1131 tggccaatgg catcggccat gtctggtcat cacgggctgg ttgcctgggt gaaggtgatg 1191 cacatcatca tgtgcggttg gaggggctgg cacagtgtgg gctgaactgg agcagttgtc 1251 caggctggcg ttgaatcagt gagggtttgt gattggcggt tgtgaagcaa tgactccgcc 1311 catattctat ttgtgggagc tgagatgatg gcatgcttgg gatgtgcatg gatcatggta 1371 gtgcagcaaa ctatattcac ctagggctgt tggtaggatc aggtgaggcc ttgcacattg 1431 catgatgtac tcgtcatggt gtgttggtga gaggatggat gtggatggat gtgtattctc Seite 68

<210> 50

<211> 322

<212> PRT

<213> Haematococcus pluvialis

<400> 50

Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile Gly 10 15

ro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu Ser 20 25 30

Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala Arg

Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser Leu 50 60

Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly Thr 65 70 75 80

Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala Leu 85 90 95

Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys Arg 100 105 110

Glu Gln Leu Ser Tyr Gln Ala Ala Ala Ile Ala Ala Ser Ile Gly Val 115 120 125

Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His Met 130 140

Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu Leu 145 150 155 160

Leu Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr Ala 165 170 175

His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His Lys 180 185 190 O4sequ.txt
Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu Phe
195 200 205 Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly Phe 210 215 220 Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu Gly 225 230 235 240 Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu Val 245 250 255 His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met Lys 260 265 270 Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly Gly 275 280 285 la Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile Pro 290 300 Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp Ser 305 310 315 320 Lys Arg <210> 51 <211> 1163 <212> DNA <213> Lycopersicon esculentum 20> 221> **CDS** <222> (1)..(942)<223> <400> 51 att cgg cac gag att tca gcc tcc gct agt tcc cga acc att cgc ctc Ile Arg His Glu Ile Ser Ala Ser Ala Ser Ser Arg Thr Ile Arg Leu 1 5 10 15 48 cgt cat aac ccg ttt ctc agt cca aaa tcc gcc tca acc gcc ccg ccg Arg His Asn Pro Phe Leu Ser Pro Lys Ser Ala Ser Thr Ala Pro Pro 20 25 30 96 gtt ctg ttc ttc tct ccg tta act cgc aat ttt ggc gca att ttg ctg Val Leu Phe Phe Ser Pro Leu Thr Arg Asn Phe Gly Ala Ile Leu Leu 144

tct Ser	aga Arg	aga Arg	aag Lys	ccg Pro	aga Arg	ttg Leu	gcg Ala	gtt Val	tgt Cys	ttt Phe	gtg Val	ctg Leu	gag Glu	aat Asn	gag Glu	192
	50					55					60					240
aaa Lys 65	Leu	aat Asn	agt Ser	Thr	atc Ile 70	gaa Glu	agt Ser	gag Glu	agt Ser	gaa Glu 75	val	Ile	Glu	Asp	Arg · 80	240
ata Ile	caa Gln	gta Val	gag Glu	att Ile 85	aat Asn	gag Glu	gag Glu	aag Lys	agt Ser 90	tta Leu	gct Ala	gcc Ala	agt Ser	tgg Trp 95	ctg Leu	288
gcg Ala	gag Glu	aaa Lys	ttg Leu 100	gcg Ala	agg Arg	aag Lys	aaa Lys	tcg Ser 105	gag Glu	agg Arg	ttt Phe	act Thr	tat Tyr 110	ctt Leu	gtg Val	336
gca Ala	gct Ala	gtg Val 115	Met	tct Ser	agt Ser	ttg Leu	ggg Gly 120	att Ile	act Thr	tct Ser	atg Met	gcg Ala 125	att Ile	ttg Leu	gcg Ala	384
gtt Val	tat Tyr 130	tac Tyr	aga Arg	ttt Phe	tca Ser	tgg Trp 135	caa Gln	atg Met	gag Glu	ggt Gly	gga Gly 140	gaa Glu	gtg Val	cct Pro	ttt Phe	432
tct Ser 145	gaa Glu	atg Met	tta Leu	gct Ala	aca Thr 150	ttc Phe	act Thr	ctc Leu	tcg Ser	ttt Phe 155	ĞĨу	gct Ala	gcc Ala	gta Val	gga Gly 160	480
															tct Ser	528
tta Leu	tgg Trp	cac His	atg Met 180	His	gag Glu	tcg Ser	cac His	cat His 185	aga Arg	cca Pro	aga Arg	gaa Glu	gga Gly 190	Pro	ttt Phe	576
gag Glu	atg Met	aac Asn 195	Asp	gtt Val	ttc Phe	gcc Ala	ata Ile 200	Thr	aat Asn	gct Ala	gtt Val	cca Pro 205	Ala	ata Ile	ggt Gly	624
ctt Leu	ctt Leu 210	Ser	tac Tyr	ggt Gly	ttc Phe	ttc Phe 215	His	aaa Lys	ggg Gly	ato Ile	gtc Val 220	Pro	ggo Gly	cto Leu	tgt LCys	672
	: Gly	gct Ala	gga Gly	ttg Leu	ggg Gly 230	Ile	aca Thr	gta Val	ttt Phe	ggg Gly 235	/ Met	g gct Ala	tao Tyi	ato Mei	ttc Phe 240	720
gtt Val	cac His	gat S Asp	gga Gly	ctg Leu 245	ı Val	cat His	aag Lys	aga Arg	ttt Phe 250	Pro	gta Val	a ggg I Gly	g cc1 / Pro	t att	t gcc e Ala 5	768
aac Asr	gtg Val	cct Pro	tac Tyr 260	Phe	cgg Arg	agg Arg	gta Val	gct Ala 265	ı Ala	a gca a Ala	a cat a His	t cag	cti Lei 27	u Hi:	t cac s His	816
tco Sei	g gad Asp	aaa D Lys 275	Phe	gat Asp	ggt Gly	gto Val	cca Pro 280	Tyı	ggd Gly	tti / Lei	g tti u Pho	t cta e Lea 28	u Gl	a cc y Pr	t aag o Lys	864
ga: Gl:	a ttg J Lei 290	u Glu	a gaa u Glu	a gta u Va	a gga l Gly	a gga / Gly 295	/ Let	gaa Glu	a gag u Gli	g tt	a gaa u Glu 30	u Ly:	g ga s Gl	a gt u Va	c aac l Asn	912
cg: Are 30	g Ar	g att	t aaa e Ly:	a att	t tct e Ser 310	Lys	g gga s Gly	a tta / Lei	a tta u Lei	a tg	atca:	aaag	ata	cgtc	tga	962

taataataaa	atgcgattgt	atttaggctg	tagattatta	ttgggaaaaa	gatagaaaga	1022
tatatatatg	aatataatat	aaaatgcaac	aagctttcta	tggagaagac	cttttcttt	1082
ttggtacctg	tacgtaaaag	gtgaacaatt	tgatgtccta	gtacttgttg	acaaaccaga	1142
agaacgataa	ttcaaaacaa	a	•			1163

<210> 52

<211> 314

<212> PRT

<213> Lycopersicon esculentum

<400> 52

Tle Arg His Glu Ile Ser Ala Ser Ala Ser Ser Arg Thr Ile Arg Leu
5 10 15

Arg His Asn Pro Phe Leu Ser Pro Lys Ser Ala Ser Thr Ala Pro Pro 20 25 30

Val Leu Phe Phe Ser Pro Leu Thr Arg Asn Phe Gly Ala Ile Leu Leu 35 40 45

Ser Arg Arg Lys Pro Arg Leu Ala Val Cys Phe Val Leu Glu Asn Glu 50 60

Lys Leu Asn Ser Thr Ile Glu Ser Glu Ser Glu Val Ile Glu Asp Arg 65 70 75 80

Ile Gln Val Glu Ile Asn Glu Glu Lys Ser Leu Ala Ala Ser Trp Leu 85 90 95

a Glu Lys Leu Ala Arg Lys Lys Ser Glu Arg Phe Thr Tyr Leu Val 100 105 110

Ala Ala Val Met Ser Ser Leu Gly Ile Thr Ser Met Ala Ile Leu Ala 115 120 125

Val Tyr Tyr Arg Phe Ser Trp Gln Met Glu Gly Gly Glu Val Pro Phe 130 140

Ser Glu Met Leu Ala Thr Phe Thr Leu Ser Phe Gly Ala Ala Val Gly 145 150 155 160

Met Glu Tyr Trp Ala Arg Trp Ala His Arg Ala Leu Trp His Ala Ser 165 170 175

Leu Trp His Met His Glu Ser His His Arg Pro Arg Glu Gly Pro Phe 180 185 190 Seite 72

Glu Met Asn Asp Val Phe Ala Ile Thr Asn Ala Val Pro Ala Ile Gly
Leu Leu Ser Tyr Gly Phe Phe His Lys Gly Ile Val Pro Gly Leu Cys
Phe Gly Ala Gly Leu Gly Ile Thr Val Phe Gly Met Ala Tyr Met Phe
225 Val His Asp Gly Leu Val His Lys Arg Phe Pro Val Gly Pro Ile Ala
Asn Val Pro Tyr Phe Arg Arg Val Ala Ala Ala His Gln Leu His His

Ser Asp Lys Phe Asp Gly Val Pro Tyr Gly Leu Phe Leu Gly Pro Lys 275 280 285

Glu Leu Glu Glu Val Gly Gly Leu Glu Glu Leu Glu Lys Glu Val Asn 290 295 300

Arg Arg Ile Lys Ile Ser Lys Gly Leu Leu 305

<210> 53

<211> 1779

<212> DNA

<213> Arabidopsis thaliana

<220>
21> CDS
<222> (1)..(1779)
<223>

atg gat ctc cgt cgg agg cct cct aaa cca ccg gtt acc aac aac aac 48

Met Asp Leu Arg Arg Arg Pro Pro Lys Pro Pro Val Thr Asn Asn Asn Asn 15

aac tcc aac gga tct ttc cgt tct tat cag cct cgc act tcc gat gac Asn Ser Asn Gly Ser Phe Arg Ser Tyr Gln Pro Arg Thr Ser Asp Asp 30

gat cat cgt cgc cgg gct aca aca att gct cct cca ccg aaa gca tcc Asp His Arg Arg Arg Ala Thr Thr Ile Ala Pro Pro Pro Lys Ala Ser 45

gac gcg ctt cct ctt ccg tta tat ctc aca aac gcc gtt ttc ttc acg 192

Seite 73

	.7.	1 011	Dro	Leu	Bro	الما	Tyr	1 60	4seq	u.tx Asn	t Ala	٧a٦	Phe	Phe	еТ	hr	
	50					55					60						240
Leu 65	Phe	Phe	Ser	gtc Val	A1a 70	Tyr	Tyr	Leu	Leu	нтs 75	Arg	ırp	arg	AS	р <u>г</u> 8	.ys 80	240
atc Ile	cgt Arg	tac Tyr	aat Asn	acg Thr 85	cct Pro	ctt Leu	cac His	gtc Val	gtc Val 90	act Thr	atc Ile	aca Thr	gaa Glu	ct Le 95	c c	igc Sly	288
gcc Ala	att Ile	att Ile	gct Ala 100	ctc Leu	atc Ile	gct Ala	tcg Ser	ttt Phe 105	atc Ile	tat Tyr	ctc Leu	cta Leu	ggg Gly 110	PII	t t	ttt Phe	336
ggt Gly	att Ile	gac Asp 115	Phe	gtt Val	cag Gln	tca Ser	ttt Phe 120	atc Ile	tca Ser	cgt Arg	gcc Ala	tct Ser 125	GIZ	ga / As	t (	gct Ala	384
tgg Trp	gat Asp 130	Leu	gcc	gat Asp	acg Thr	atc Ile 135	gat Asp	gat Asp	gat Asp	gac Asp	cac His 140	Arg	cti Lei	t gt u Va	ic :	acg Thr	432
ys 145	ser	cca	ccg Pro	act Thr	ccg Pro 150	Ile	gtt Val	tcc Ser	gtt Val	gct Ala 155	Lys	tta Lei	ı cc	t aa o As	at sn	ccg Pro 160	480
gaa Glu	cct Pro	att Ile	gtt Val	acc Thr 165	Glu	tcg Ser	ctt Leu	cct	gag Glu 170	Glu	gac Asp	gag Gli	g ga u Gl	u i	tt le 75	gtg Val	528
aaa Lys	tcg S Sei	g gti Va	t ato 1 Ilo 180	c gad e Asp O	gga Gly	gtt Val	att Ile	cca Pro 185	Ser	tac Tyr	tcg Sei	g ct	t ga u Gl 19	u >	ct er	cgt Arg	576
cto	n GJ	t ga y As 19	p Cy:	c aaa s Ly:	a aga s Arg	a gcg g Ala	g gcg a Ala 200	a Sei	att Ile	cgt Arg	g Ar	t ga g G1 20	u Ai	g t a L	tg .eu	cag Gln	624
ag: Ar	a gt g Va 21	] Th	c gg r Gl	g ag y Ar	a tcg g Sei	g ati r Ilo 21	e GII	a ggg u Gly	g tta y Lei	a cci	g tte Le 22	ų AS	t go p G	a t y P	tt he	gat Asp	672
ta Ty 2	t ga r Gl 5	a tc u Se	g at r Il	t tt e Le	g gg u G1 23	y Gli	a tg n Cy:	c tg s Cy	t gag s Gli	g ate u Me 23	r Pr	t gt o Va	t gg	ga t ly T	ac Tyr	att Ile 240	720
Za Gl	g at n Il	t cc e Pr	t gt o Va	t gg 1 G1 24	УII	t gc e Al	t gg a Gl	t cc y Pr	a ttg o Leg 25	u Le	g ct u Le	t ga u As	t g sp G	ıy :	tat Tyr 255	gag Glu	768
ta Ty	c to r se	t gt er Va	t co il Pr 26	o Me	g gc t Al	t ac a Th	a ac r Th	c ga r Gl 26	u GI	t tg y Cy	t tt 's Le	g gt eu Va	A LE	ct a la : 70	agc Ser	act Thr	816
aa As	ic ag in Ar	g G	jc to ly cy 75	gc aa /s Ly	ig go /s Al	t at a Me	g tt t Ph 28	ie II	c tc e Se	t gg r Gl	t gg y G	IY A	cc a la T 85	cc hr	agt Ser	acc Thr	864
gt Va	al Le	ct aa eu Ly 90	ag ga ys A:	ac go sp G	gt at ly Me	g acet Th	ır Ar	ga go 'g Al	a co la Pr	t gt o Va	II V	tt c al A DO	gg t rg P	tc he	gct Ala	t tcg a Ser	912
A	cg ag la Ai 05	ga c rg A	ga g rg A	ct to la Se	er G	ag ct lu Le 10	t aa eu Ly	ag ti /s Pl	it tt 1e Ph	ie Le	tg g eu G 15	ag a lu A	at c sn F	ca Pro	gaç Gli	g aac u Asn 320	960
t <sup>.</sup>	tt g	at a	ct t	tg g	ca g	ta g	tciti	tc a		gg to eite		gt a	ga t	tt	gc	a aga	1008

# O4sequ.txt Phe Asp Thr Leu Ala Val Val Phe Asn Arg Ser Ser Arg Phe Ala Arg 325 330 335

ctg caa agt gtt aaa tgc aca atc gcg ggg aag aat gct tat gta agg Leu Gln Ser Val Lys Cys Thr Ile Ala Gly Lys Asn Ala Tyr Val Arg 340 . 345 350

ttc tgt tgt agt act ggt gat gct atg ggg atg aat atg gtt tct aaa Phe Cys Cys Ser Thr Gly Asp Ala Met Gly Met Asn Met Val Ser Lys 355 360

1056

ggt gtg cag aat gtt ctt gag tat ctt acc gat gat ttc cct gac atg
Gly Val Gln Asn Val Leu Glu Tyr Leu Thr Asp Asp Phe Pro Asp Met
370
375
380

gat gtg att gga atc tct ggt aac ttc tgt tcg gac aag aaa cct gct Asp Val Ile Gly Ile Ser Gly Asn Phe Cys Ser Asp Lys Lys Pro Ala 385 390 395 400

gct gtg aac tgg att gag gga cgt ggt aaa tca gtt gtt tgc gag gct 1248 Ala Val Asn Trp Ile Glu Gly Arg Gly Lys Ser Val Val Cys Glu Ala 405 410 415

ta atc aga gga gag atc gtg aac aag gtc ttg aaa acg agc gtg gct al Ile Arg Gly Glu Ile Val Asn Lys Val Leu Lys Thr Ser Val Ala 420 425 430

gct tta gtc gag ctc aac atg ctc aag aac cta gct ggc tct gct gtt Ala Leu Val Glu Leu Asn Met Leu Lys Asn Leu Ala Gly Ser Ala Val
435
440
445

gca ggc tct cta ggt gga ttc aac gct cat gcc agt aac ata gtg tct
Ala Gly Ser Leu Gly Gly Phe Asn Ala His Ala Ser Asn Ile Val Ser
450
450
460

gct gta ttc ata gct act ggc caa gat cca gct caa aac gtg gag agt
Ala Val Phe Ile Ala Thr Gly Gln Asp Pro Ala Gln Asn Val Glu Ser
465 470 475 480

tct caa tgc atc acc atg atg gaa gct att aat gac ggc aaa gat atc Ser Gln Cys Ile Thr Met Met Glu Ala Ile Asn Asp Gly Lys Asp Ile 485 490 495

cat atc tca gtc act atg cca tct atc gag gtg ggg aca gtg gga gga 1536 His Ile Ser Val Thr Met Pro Ser Ile Glu Val Gly Thr Val Gly Gly 500 505 510

ga aca cag ctt gca tct caa tca gcg tgt tta aac ctg ctc gga gtt Gly Thr Gln Leu Ala Ser Gln Ser Ala Cys Leu Asn Leu Leu Gly Val 515 520

aaa gga gca agc aca gag tcg ccg gga atg aac gca agg agg cta gcg
Lys Gly Ala Ser Thr Glu Ser Pro Gly Met Asn Ala Arg Arg Leu Ala
530
535
540

acg atc gta gcc gga gca gtt tta gct gga gag tta tct tta atg tca
Thr Ile Val Ala Gly Ala Val Leu Ala Gly Glu Leu Ser Leu Met Ser
545 550 555 560

gca att gca gct gga cag ctt gtg aga agt cac atg aaa tac aat aga 1728 Ala Ile Ala Ala Gly Gln Leu Val Arg Ser His Met Lys Tyr Asn Arg 565 570 575

tga 1779

```
<210> 54
```

<211> 592

<212> PRT

<213> Arabidopsis thaliana

<400> 54

Met Asp Leu Arg Arg Pro Pro Lys Pro Pro Val Thr Asn Asn Asn 1 10 15

Asn Ser Asn Gly Ser Phe Arg Ser Tyr Gln Pro Arg Thr Ser Asp Asp 20 25 30

sp His Arg Arg Arg Ala Thr Thr Ile Ala Pro Pro Pro Lys Ala Ser 35 40 45

Asp Ala Leu Pro Leu Pro Leu Tyr Leu Thr Asn Ala Val Phe Phe Thr 50 60

Leu Phe Phe Ser Val Ala Tyr Tyr Leu Leu His Arg Trp Arg Asp Lys 70 75 80

Ile Arg Tyr Asn Thr Pro Leu His Val Val Thr Ile Thr Glu Leu Gly 85 90 95

Ala Ile Ile Ala Leu Ile Ala Ser Phe Ile Tyr Leu Leu Gly Phe Phe 100 105 110

Gly Ile Asp Phe Val Gln Ser Phe Ile Ser Arg Ala Ser Gly Asp Ala 115 120 125

p Asp Leu Ala Asp Thr Ile Asp Asp Asp Asp His Arg Leu Val Thr 130 140

Cys Ser Pro Pro Thr Pro Ile Val Ser Val Ala Lys Leu Pro Asn Pro 145 150 160

Glu Pro Ile Val Thr Glu Ser Leu Pro Glu Glu Asp Glu Glu Ile Val 165 170 175

Lys Ser Val Ile Asp Gly Val Ile Pro Ser Tyr Ser Leu Glu Ser Arg 180 185 190

Leu Gly Asp Cys Lys Arg Ala Ala Ser Ile Arg Arg Glu Ala Leu Gln 195 200 205

Arg Val Thr Gly Arg Ser Ile Glu Gly Leu Pro Leu Asp Gly Phe Asp 210 220 Seite 76

Tyr Glu Ser Ile Leu Gly Gln Cys Cys Glu Met Pro Val Gly Tyr Ile 225 230 235 240 Gln Ile Pro Val Gly Ile Ala Gly Pro Leu Leu Asp Gly Tyr Glu 245 250 255 Tyr Ser Val Pro Met Ala Thr Thr Glu Gly Cys Leu Val Ala Ser Thr 260 265 270 Asn Arg Gly Cys Lys Ala Met Phe Ile Ser Gly Gly Ala Thr Ser Thr 275 280 285 Val Leu Lys Asp Gly Met Thr Arg Ala Pro Val Val Arg Phe Ala Ser 290 295 300 la Arg Arg Ala Ser Glu Leu Lys Phe Phe Leu Glu Asn Pro Glu Asn 310 315 320 Phe Asp Thr Leu Ala Val Val Phe Asn Arg Ser Ser Arg Phe Ala Arg 325 330 335 Leu Gln Ser Val Lys Cys Thr Ile Ala Gly Lys Asn Ala Tyr Val Arg 340 345 350 Phe Cys Cys Ser Thr Gly Asp Ala Met Gly Met Asn Met Val Ser Lys 355 360 Gly Val Gln Asn Val Leu Glu Tyr Leu Thr Asp Asp Phe Pro Asp Met 370 380 Asp Val Ile Gly Ile Ser Gly Asn Phe Cys Ser Asp Lys Lys Pro Ala 385 390 395 a Val Asn Trp Ile Glu Gly Arg Gly Lys Ser Val Val Cys Glu Ala 405 415 Val Ile Arg Gly Glu Ile Val Asn Lys Val Leu Lys Thr Ser Val Ala 420 425 430 Ala Leu Val Glu Leu Asn Met Leu Lys Asn Leu Ala Gly Ser Ala Val 435 440 445 Ala Gly Ser Leu Gly Gly Phe Asn Ala His Ala Ser Asn Ile Val Ser 450 460 Ala Val Phe Ile Ala Thr Gly Gln Asp Pro Ala Gln Asn Val Glu Ser 465 470 475 480 Ser Gln Cys Ile Thr Met Met Glu Ala Ile Asn Asp Gly Lys Asp Ile 485 490 495

His	Ile	Ser	va1 500	Thr	Met	Pro	Ser	Ile <sup>-</sup> 505	Glu	٧al	Gly	Thr	va1 510	Glу	Gly
Gly	Thr	Gln 515	Leu	Αla	Ser	Gln	ser 520	Аlа	Cys	Leu	Asn	Leu 525	Leu	Glу	val
Lys	Gly 530		Ser	Thr	Glu	Ser 535	Pro	Gly	Met	Asn	Ala 540	Arg	Arg	Leu	Αla
Thr 545		۷al	Ala	GТу	Ala 550	Val	Leu	Αla	Gly	Glu 555	Leu	Ser	Leu	Met	Ser 560
۸la	Ile	Αla	Ala	Gly 565	Gln	Leu	'Val	Arg	Ser 570	His	Met	Lys	Tyr	Asn 575	Arg
er	ser	Arg	Asp	Ile	Ser	GТу	Αla	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Thr

580 585 590

<210> 55

<211> 1401

<212> DNA

<213> Arabidopsis thaliana ISPH

<220>

<221> CDS

<222> (1)..(1401)

<223>

ta	)> 5 gct Ala	att	gcg Ala	ctc Leu 5	caa Gln	ttc Phe	agc Ser	cga Arg	tta Leu 10	tgc Cys	gtt Val	cga Arg	ccg Pro	gat Asp 15	act Thr		48
ttc Phe	gtg Val	cgg Arg	gag Glu 20	aat Asn	cat His	ctc Leu	tct Ser	gga Gly 25	tcc Ser	gga Gly	tct Ser	ctc Leu	cgc Arg 30	cgc Arg	cgg Arg		96
aaa Lys	gct Ala	tta Leu 35	tca Ser	gtc Val	cgg Arg	tgc Cys	tcg Ser 40	tct Ser	ggc Gly	gat Asp	gag Glu	aac Asn 45	gct Ala	cct Pro	tcg Ser	:	144
cca Pro	tcg Ser 50	gtg Val	gtg Val	atg Met	gac Asp	tcc Ser 55	gat Asp	ttc Phe	gac Asp	gcc Ala	aag Lys 60	gtg Val	ttc Phe	cgt Arg	aag Lys	:	192
aac Asn 65	ttg Leu	acg Thr	aga Arg	agc Ser	gat Asp 70	aat Asn	tac Tyr	aat Asn	cgt Arg	aaa Lys 75	ggg Gly	ttc Phe	ggt Gly	cat His	aag Lys 80		240
gag	gag	aca	ctc	aag	ctc	atg	aat	cga ·		tac		agt	gat	ata	ttg		288

Glu Glu Thr Leu Lys Leu Met Asn Arg Glu Tyr Thr Ser Asp Ile Leu 85 90 95 gag aca ctg aaa aca aat ggg tat act tat tct tgg gga gat gtt act Glu Thr Leu Lys Thr Asn Gly Tyr Thr Tyr Ser Trp Gly Asp Val Thr 100 105 110336 gtg aaa ctc gct aaa gca tat ggt ttt tgc tgg ggt gtt gag cgt gct Val Lys Leu Ala Lys Ala Tyr Gly Phe Cys Trp Gly Val Glu Arg Ala 115 120 125 384 gtt cag att gca tat gaa gca cga aag cag ttt cca gag gag agg ctt Val Gln Ile Ala Tyr Glu Ala Arg Lys Gln Phe Pro Glu Glu Arg Leu 130 135 140 432 tgg att act aac gaa atc att cat aac ccg acc gtc aat aag agg ttg Trp Ile Thr Asn Glu Ile Ile His Asn Pro Thr Val Asn Lys Arg Leu 145 150 155 160 480 gaa gat atg gat gtt aaa att att ccg gtt gag gat tca aag aaa cag Glu Asp Met Asp Val Lys Ile Ile Pro Val Glu Asp Ser Lys Lys Gln 165 170 175 528 t gat gta gta gag aaa gat gat gtg gtt atc ctt cct gcg ttt gga he Asp Val Val Glu Lys Asp Asp Val Val Ile Leu Pro Ala Phe Gly 180 185 576 gct ggt gtt gac gag atg tat gtt ctt aat gat aaa aag gtg caa att Ala Gly Val Asp Glu Met Tyr Val Leu Asn Asp Lys Lys Val Gln Ile 195 200 205 624 gtt gac acg act tgt cct tgg gtg aca aag gtc tgg aac acg gtt gag Val Asp Thr Thr Cys Pro Trp Val Thr Lys Val Trp Asn Thr Val Glu 210 215 220 672 aag cac aag aag ggg gaa tac aca tca gta atc cat ggt aaa tat aat Lys His Lys Lys Gly Glu Tyr Thr Ser Val Ile His Gly Lys Tyr Asn 225 230 240 720 cat gaa gag acg att gca act gcg tct ttt gca gga aag tac atc att His Glu Glu Thr Ile Ala Thr Ala Ser Phe Ala Gly Lys Tyr Ile Ile 245 250 255 768 gta aag aac atg aaa gag gca aat tac gtt tgt gat tac att ctc ggt Val Lys Asn Met Lys Glu Ala Asn Tyr Val Cys Asp Tyr Ile Leu Gly 260 265 270 816 gc caa tac gat gga tct agc tcc aca aaa gag gag ttc atg gag aaa Gly Gln Tyr Asp Gly Ser Ser Ser Thr Lys Glu Glu Phe Met Glu Lys 275 280 285 864 ttc aaa tac gca att tcg aag ggt ttc gat ccc gac aat gac ctt gtc Phe Lys Tyr Ala Ile Ser Lys Gly Phe Asp Pro Asp Asn Asp Leu Val 290 295 300 912 aaa gtt ggt att gca aac caa aca acg atg cta aag gga gaa aca gag Lys Val Gly Ile Ala Asn Gln Thr Thr Met Leu Lys Gly Glu Thr Glu 305 310 315 960 gag ata gga aga tta ctc gag aca aca atg atg cgc aag tat gga gtg Glu Ile Gly Arg Leu Leu Glu Thr Thr Met Met Arg Lys Tyr Gly Val 325 330 335 1008

gaa aat gta agc gga cat ttc atc agc ttc aac aca ata tgc gac gct Glu Asn Val Ser Gly His Phe Ile Ser Phe Asn Thr Ile Cys Asp Ala 340

act caa gag cga caa gac gca atc tat gag cta gtg gaa gag aag att Seite 79

Thr	Gln		Arg	Gln	Asp	Ala		Tyr	)4sec Glu	u.tx Leu	t Val	Glu	Glu	Lys	Ile	
gac Asp	ctc Leu	355 atg Met	cta Leu	gtg Val	gtt Val	ggc Gly	360 gga Gly	tgg Trp	aat Asn	tca <sup>°</sup> Ser	agt Ser	aac Asn	acc Thr	tct Ser	cac His	1152
	370					375 gca					380					1200
Leu 385	ĞÎñ	ĞTü	Ile	Ser	ัฐโน 390	Ăla	Arg	Ğโy	Ile	Pro 395	Ser	Tyr	Třp	Ile	Asp 400	
agt Ser	gag Glu	aaa Lys	cgg Arg	ata Ile 405	gga Gly	cct Pro	ggg Gly	aat Asn	aaa Lys 410	ata Ile	gcc Ala	tat Tyr	aag Lys	ctc Leu 415	cac His	1248
tat Tyr	gga Gly	gaa Glu	ctg Leu 420	gtc Val	gag Glu	aag Lys	gaa Glu	aac Asn 425	ttt Phe	ctc Leu	cca Pro	aag Lys	gga Gly 430	cca Pro	ata Ile	1296
aca Thr	atc Ile	ggt Gly 435	gtg Val	aca Thr	tca Ser	ggt Gly	gca Ala 440	tca Ser	acc Thr	ccg Pro	gat Asp	aag Lys 445	gtc Val	gtg Val	gaa Glu	1344
it sp	gct Ala 450	Leu	gtg Val	aag Lys	gtg Val	ttc Phe 455	gac Asp	att Ile	aaa Lys	cgt Arg	gaa Glu 460	gag Glu	tta Leu	ttg Leu	cag Gln	1392
	Ā٦a	tga				٠										1401
<21	0>	56														
<21	1>	466			٠											
<21	2>	PRT														
<21	3>	Arab	oi dop	sis	thal	iana	ISP	Н								
<40	0>	56														
Met	: Ala	ı Val	Аla	Leu 5	G]r	) Phe	ser	Arg	Leu 10	Cys	val	Arg	Pro	15	Thr	
Phe	· Val	i Arg	g Glu 20	ı Asn	Hi≤	Leu	Ser	Gly 25	' Ser	· Gly	' Ser	Leu	arg 30	y Arg	, Arg	
Lys	a Ta	Leu 35	ı Ser	· val	l Arg	g Cys	Ser 40	· Ser	· Gly	/ Asp	G]ı	ı Asr 45	ı. Ala	a Pro	ser	
Pro	Sei 50	r Va	l val	l Met	: As	Ser 55	. Ast	Phe	e Asp	Ala	60	s Val	Pho	e Ar	g Lys	
Asr 65	ı Lei	u Thi	r Arg	g Sei	70	o Asr	туі	- Asr	n Ar	g Lys 75	s Gly	y Phe	e Gly	y Hi:	s Lys 80	
Glı	ı Gl	u Th	r Lei	и Lys 85	s Lei	u Me1	t Ası	n Arg	g G7i 90	и Туі	r Th	r Sei	^ As	p Il 95	e Leu	

Glu Thr Leu Lys Thr Asn Gly Tyr Thr Tyr Ser Trp Gly Asp Val Thr 100 105 110 Val Lys Leu Ala Lys Ala Tyr Gly Phe Cys Trp Gly Val Glu Arg Ala 115 120 125 Val Gln Ile Ala Tyr Glu Ala Arg Lys Gln Phe Pro Glu Glu Arg Leu 130 140 Trp Ile Thr Asn Glu Ile Ile His Asn Pro Thr Val Asn Lys Arg Leu 145 155 160 Glu Asp Met Asp Val Lys Ile Ile Pro Val Glu Asp Ser Lys Lys Gln 165 170 175 Phe Asp Val Val Glu Lys Asp Asp Val Val Ile Leu Pro Ala Phe Gly 180 185 190 ia Gly Val Asp Glu Met Tyr Val Leu Asn Asp Lys Lys Val Gln Ile 195 200 205 Val Asp Thr Thr Cys Pro Trp Val Thr Lys Val Trp Asn Thr Val Glu 210 215 220 Lys His Lys Lys Gly Glu Tyr Thr Ser Val Ile His Gly Lys Tyr Asn 235 240 His Glu Glu Thr Ile Ala Thr Ala Ser Phe Ala Gly Lys Tyr Ile Ile 245 250 255 Val Lys Asn Met Lys Glu Ala Asn Tyr Val Cys Asp Tyr Ile Leu Gly 260 265 270 Gly Gln Tyr Asp Gly Ser Ser Ser Thr Lys Glu Glu Phe Met Glu Lys 275 280 285 Phe Lys Tyr Ala Ile Ser Lys Gly Phe Asp Pro Asp Asn Asp Leu Val 290 295 300 Lys Val Gly Ile Ala Asn Gln Thr Thr Met Leu Lys Gly Glu Thr Glu 305 310 315 320 Glu Ile Gly Arg Leu Leu Glu Thr Thr Met Met Arg Lys Tyr Gly Val 325 330 335 Glu Asn Val Ser Gly His Phe Ile Ser Phe Asn Thr Ile Cys Asp Ala 340 345 350 Thr Gln Glu Arg Gln Asp Ala Ile Tyr Glu Leu Val Glu Glu Lys Ile 355 360 365

Asp Leu Met Leu Val Val Gly Trp Asn Ser Ser Asn Thr Ser His 370 Met Leu Val Val Gly Trp Asn Ser Ser Ser Asn Thr Ser His 385 Asn Thr Ser Gly Ala Arg Gly Ile Pro Ser Tyr Trp Ile Asp 400 Asn Lys Ile Ala Tyr Lys Leu His 410 Tyr Gly Glu Leu Val Glu Lys Glu Asn Phe Leu Pro Lys Gly Pro Ile 420 Asn Asp Ala Leu Val Lys Val Phe Asp Ile Lys Arg Glu Glu Leu Leu Gln 455 Asn Ala Leu Val Lys Val Phe Asp Ile Lys Arg Glu Glu Leu Leu Gln 455 DNA Ala Lycopersicon esculentum  4210 S7  4211 Z160  4212 DNA  4220 C222 (1)(2160)  4223 DNA  400 S7  400 S8  400 S7  400 S8  40	
Ser Glu Lys Arg Ile Gly Pro Gly Asn Lys Ile Ala Tyr Lys Leu His 415  Tyr Gly Glu Leu Val Glu Lys Glu Asn Phe Leu Pro Lys Gly Pro Ile 420  Thr Ile Gly Val Thr Ser Gly Ala Ser Thr Pro Asp Lys Val Val Glu Asp Ala Leu Val Lys Val Phe Asp Ile Lys Arg Glu Glu Leu Leu Gln 450  Asp Ala Leu Val Lys Val Phe Asp Ile Lys Arg Glu Glu Leu Leu Gln 450  Au Ala 65  <210> 57  <211> 2160  <212> DNA  <213> Lycopersicon esculentum  <220> <221> CDS  <222> (1)(2160)  <223>  400> 57  atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly 15  gtg gtt tca gat tct tct aag gca acc cct ttg ttc tct gga tgg att Val	
Tyr Gly Glu Leu Val Glu Lys Glu Asn Phe Leu Pro Lys Gly Pro Ile 420  Thr Ile Gly Val Thr Ser Gly Ala Ser Thr Pro Asp Lys Val Val Glu 435  Asp Ala Leu Val Lys Val Phe Asp Ile Lys Arg Glu Glu Leu Leu Gln 455  Asp Ala Leu Val Lys Val Phe Asp Ile Lys Arg Glu Glu Leu Leu Gln 455  400  57  <211> 2160  <212> DNA  <213> Lycopersicon esculentum  <220> <221> CDS  <222> (1)(2160)  <223>  400> 57  atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly 10  gtg gtt tca gat tct tct aag gca acc cct ttg ttc tct gga tgg att Val Val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile 20  30  4103  415  415  416  417  418  418  418  418  419  419  419  419	
Thr Ile Gly Val Thr Ser Gly Ala Ser Thr Pro Asp Lys Val Val Glu  Asp Ala Leu Val Lys Val Phe Asp Ile Lys Arg Glu Glu Leu Leu Gln  au Ala  65  <210> 57  <211> 2160  <212> DNA  <213> Lycopersicon esculentum	
Asp Ala Leu Val Lys Val Phe Asp Ile Lys Arg Glu Glu Leu Leu Gln  Au Ala  210> 57  2210> 57  2211> 2160  2212> DNA  2213> Lycopersicon esculentum   2220>  2221> CDS  2222> (1)(2160)  2223>  Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly  1  gtg gtt tca gat tct tct aag gca acc cct ttg ttc tct gga tgg att  Yal Val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile  20  Asp Ala Leu Cys Ala Thr Pro Leu Phe Ser Gly Trp Ile  2405  250  Ala Tyr Ala Thr Pro Leu Phe Ser Gly Trp Ile  250  Asp Ala Tyr Ala Thr Pro Leu Phe Ser Gly Trp Ile  260  Asp Ala Tyr Ala Thr Pro Leu Phe Ser Gly Trp Ile  270  Ala Thr Pro Leu Phe Ser Gly Trp Ile	
au Ala 65 Ala  <210> 57  <211> 2160  <212> DNA  <213> Lycopersicon esculentum  <220> <221> CDS  <222> (1)(2160)  <223>  400> 57  atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly 1  gtg gtt tca gat tct tct aag gca acc cct ttg ttc tct gga tgg att Val Val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile 20  210  221  225  220  221  225  220  230  2400	
<pre> &lt;210&gt; 57 &lt;211&gt; 2160 &lt;212&gt; DNA &lt;213&gt; Lycopersicon esculentum  &lt;220&gt; &lt;221&gt; CDS &lt;222&gt; (1)(2160) &lt;223&gt;  400&gt; 57 atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly 15 gtg gtt tca gat tct tct aag gca acc cct ttg ttc tct gga ttg att val val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile 20</pre>	
<pre>&lt;211&gt; 2160 &lt;212&gt; DNA &lt;213&gt; Lycopersicon esculentum  </pre> <pre>&lt;220&gt; &lt;221&gt; CDS &lt;222&gt; (1)(2160) &lt;223&gt;  400&gt; 57 atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly 1 gtg gtt tca gat tct tct aag gca acc cct ttg ttc tct gga ttg att Val Val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile 20</pre>	
<pre>&lt;212&gt; DNA &lt;213&gt; Lycopersicon esculentum  &lt;220&gt; &lt;221&gt; CDS &lt;222&gt; (1)(2160) &lt;223&gt;  400&gt; 57 atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly 1 gtg gtt tca gat tct tct aag gca acc cct ttg ttc tct gga tgg att Val Val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile 20</pre>	
<pre>&lt;213&gt; Lycopersicon esculentum  &lt;220&gt; &lt;221&gt; CDS &lt;222&gt; (1)(2160) &lt;223&gt;  400&gt; 57 atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly 1</pre>	
<pre>&lt;220&gt; &lt;221&gt; CDS &lt;222&gt; (1)(2160) &lt;223&gt;  400&gt; 57 atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly 1 gtg gtt tca gat tct tct aag gca acc cct ttg ttc tct gga tgg att Val Val Ser Asp ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile 20</pre>	
<pre>&lt;221&gt; CDS &lt;222&gt; (1)(2160) &lt;223&gt;  400&gt; 57 atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly 1 gtg gtt tca gat tct tct aag gca acc cct ttg ttc tct gga tgg att Val Val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile 20</pre>	
<pre>&lt;222&gt; (1)(2160) &lt;223&gt;  400&gt; 57 atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly 1 gtg gtt tca gat tct tct aag gca acc cct ttg ttc tct gga tgg att Val Val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile 20</pre>	
<pre>&lt;223&gt; 400&gt; 57 atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly 1 gtg gtt tca gat tct tct aag gca acc cct ttg ttc tct gga tgg att Val Val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile 20 25</pre>	
atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly 1 10 15  gtg gtt tca gat tct tct aag gca acc cct ttg ttc tct gga tgg att Val Val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile 20 25	
atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly 1	
atg gct ttg tgt gct tat gca ttt cct ggg att ttg aac agg act ggt Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly 1	
20 25 25 30	48
cat gga aca gat ctg cag ttt ttg ttc caa cac aag ctt act cat gag His Gly Thr Asp Leu Gln Phe Leu Phe Gln His Lys Leu Thr His Glu 35 40 45	96
gtc aag aaa agg tca cgt gtg gtt cag gct tcc tta tca gaa tct gga Val Lys Lys Arg Ser Arg Val Val Gln Ala Ser Leu Ser Glu Ser Gly 50 55 60	96 144
gaa tac tac aca cag aga ccg cca acg cct att ttg gac act gtg aac Glu Tyr Tyr Thr Gln Arg Pro Pro Thr Pro Ile Leu Asp Thr Val Asn 75 80	

tat Tyr	ccc Pro	att Ile	cat His	atg Met 85	aaa Lys	aat Asn	ctg Leu	tct Ser	ctg Leu 90	aag Lys	gaa Glu	ctt Leu	aaa Lys	caa Gln 95	cta Leu	288
gca Ala	gat Asp	gaa Glu	cta Leu 100	agg Arg	tca Ser	gat Asp	aca Thr	att Ile 105	ttc Phe	aat Asn	gta Val	tca Ser	aag Lys 110	act Thr	ggg Gly	336
ggt Gly	cac His	ctt Leu 115	ggc Gly	tca Ser	agt Ser	ctt Leu	ggt Gly 120	gtt Val	gtt Val	gag Glu	ctg Leu	act Thr 125	gtt Val	gct Ala	ctt Leu	384
cat His	tat Tyr 130	gtc Val	ttc Phe	aat Asn	gca Ala	ccg Pro 135	caa Gln	gat Asp	agg Arg	att Ile	ctc Leu 140	tgg Trp	gat Asp	gtt Val	ggt Gly	432
cat His 145	cag Gln	tct Ser	tat Tyr	cct Pro	cac His 150	aaa Lys	atc Ile	ttg Leu	act Thr	ggt Gly 155	Arg	agg Arg	gac Asp	aag Lys	atg Met 160	480
tcg ser	aca Thr	tta Leu	agg Arg	cag Gln 165	aca Thr	gat Asp	ggt Gly	ctt Leu	gca Ala 170	gga Gly	ttt Phe	act Thr	aag Lys	cga Arg 175	tcg Ser	528
ag Glu	agt Ser	gaa Glu	tat Tyr 180	Asp	tgc Cys	ttt Phe	ggc Gly	acc Thr 185	ggc Gly	cac His	agt Ser	tcc Ser	acc Thr 190	Thi	atc lle	576
tca Ser	gca Ala	ggc Gly 195	cta Leu	ggg Gly	atg Met	gct Ạla	gtt Val 200	Gly	aga Arg	gat Asp	cta Leu	aaa Lys 205	Gly	aga Ar	a aac g Asn	624
aac Asn	aat Asn 210	٧a٦	att Ile	gcc Ala	gta Val	ata Ile 215	Gly	gat Asp	ggt Gly	gco Ala	atg Met 220	Thi	a gca r Ala	a gg	t caa y Gln	672
gct Ala 225	Tyr	gaa Glu	gcc	atg Met	aat Asn 230	Asr	gct Ala	ggt Gly	tac Tyr	cto Lei 23	i Yel	tc Se	t gao r Ası	at Me	g att t Ile 240	
gtt Val	ato	tta Leu	aac Asr	gac n Asp 245	) Asn	aga Arg	caa g Glr	a gtt n Val	tct Ser 250	. Lei	a cc u Pr	t ac o Th	t gc r Ala	t ac a Th 25	t ctg r Leu 5	768
gat	ggg Gly	CCa Pro	gt <u>1</u> Va 260	l Ala	cct a Pro	gt <u>i</u> Va	gga l Gly	a gct y Ala 265	<u>a</u> Lei	ı se a ag	t ag r Se	t gc r Al	t tt a Le 27	u Se	c agg r Arg	816 I
tta Leu	cag Glr	tct Ser 275	ASI	t agg	g cct g Pro	cto Le	aga u Ar 28	g Glu	a cta u Lei	a ag u Ar	a ga g Gl	a gt u Va 28	I At	a aa a Ly	ig gga 's Gly	864 '
gtt Va	act Thi 290	Lys	g ca s Gl	g ati n Ilo	t ggt e Gly	t gg y G1: 29	y Pr	t ate	g ca t Hi	t ga s Gl	g ct u Le 30	u Al	t go a Al	a aa a Ly	a gtt ⁄s Val	912 I
ga1 Ast 30!	o Glu	a ta u Ty	t gc r Al	t cg a Ar	t gg g Gly 31	y Me	g at t Il	t ag e Se	t gg r Gl	t to y Se 31	r GI	a to y Se	a ac er Th	a ti ir Le	g tti eu Phe 320	e
ga: Gl:	a gaa u Gl	a ct <sup>.</sup>	t gg u Gl	a ct y Le 32	цТу	c ta r Ty	t at r Il	t gg e Gl	t cc y Pr 33	o Va	g ga ll As	it go sp G	gt ca ly Hi	S A	ac ati sn Ilo 35	t 1008 e
ga As	t ga p As	t ct p Le	a at u Il 34	e Al	g at a Il	t ct e Le	c aa u Ly	a ga 's Gl 34	u Va ·5	t ag l Ar	rg Se	gt ac er Ti	ir Ly	aa a /s T 50	ca aca hr Th	a 1056 r

	ggt ( Gly	Pro '	gta Val 355	ctg Leu	atc Ile	cat His	gtt Val	gtc Val 360	act Thr	gag Glu	aaa Lys	ggc Gly	aga Arg 365	ggt Gly	tat Tyr	cca Pro	1104
	tat Tyr	gct Ala 370	gag Glu	aga Arg	gct Ala	gca Ala	gat Asp 375	aag Lys	tat Tyr	cat His	gga Gly	gtt Val 380	gcc Ala	aag Lys	ttt Phe	gat Asp	1152
	cca Pro 385	gca Ala	aca Thr	gga Gly	aag Lys	caa Gln 390	ttc Phe	aaa Lys	gcc Ala	agt Ser	gcc Ala 395	aag Lys	aca Thr	cag Gln	tcc Ser	tat Tyr 400	1200
	aca Thr	aca Thr	tat Tyr	ttt Phe	gcc Ala 405	gag Glu	gct Ala	tta Leu	att Ile	gca Ala 410	gaa Glu	gca Ala	gaa Glu	gca Ala	gat Asp 415	Lys	1248
	gac Asp	att Ile	gtt val	gca Ala 420	atc Ile	cat His	gct Ala	gcc Ala	atg Met 425	ggg Gly	ggt Gly	ggg Gly	acc Thr	gga Gly 430	ME	aac Asn	1296
	ctt	ttc Phe	cat His 435	Arg	cgc Arg	ttc Phe	cca Pro	aca Thr 440	agg Arg	tgt Cys	ttt Phe	gat Asp	gtt Val 445	Gly	ata Ile	gca Ala	1344
	gaa Glu	caa Gln 450	His	gca Ala	gta Val	acc Thr	ttt Phe 455	gct Ala	gct Ala	gga Gly	ttg Leu	gct Ala 460	Cys	gaa Glu	ggo Gly	att / Ile	1392
	aaa Lys 465	cct Pro	ttc Phe	tgt Cys	gca Ala	atc Ile 470	Tyr	tcg Ser	tct Ser	ttc Phe	atg Met 475	GIT	agg Arg	g gct g Ala	ta a Ty	t gac r Asp 480	1440
	cag Gln	gta Val	gtg Val	cat His	gac Asp 485	Val	gat Asp	ttg Lei	caa Gln	aag Lys 490	Leu	CCC Pro	gto Va	g agg	g tt g Ph 49	t gca e Ala 5	1488
	atg Met	gac Asp	aga Arg	g gca g Ala 500	Gly	ctt Leu	gtt Val	gga Gly	a gca / Ala 505	<u>a</u> Asp	ggt Gly	CCa Pro	a aca	a ca r Hi 51	s cy	t ggt s Gly	1536
	gca Ala	ttt Phe	ga Ası 51	o Va	t act I Thr	tad Tyr	ato Met	g gca 5 Ala 520	a Cys	t ctt s Lei	cct Pro	t aa O As	c at n Me 52	t va	t gt 1 Va	a atg 1 Met	1584 :
	gct	cct Pro 530	Se	t ga r As	t gaa p Glu	a gcg a Ala	g gag a Gli 53!	ı Lei	a tti u Pho	t cad e His	ate Me	g gt t Va 54	IAI	a ac a Th	t go r Al	t gcc a Ala	1632 1
	gco A1a 545	11e	t ga e As	t ga p As	c aga p Ar	a cca g Pro 550	o Se	t tg r Cy	t tt s Ph	t aga e Ar	a ta g Ty 55	r Pr	a ag o Ar	a gg g Gl	ya aa y As	it ggg sn Gly 560	<u> </u>
	ato Ilo	c gg e Gl	t gt y Va	a ga 1 G1	g ct <sup>.</sup> u Le 56	u Pr	g gc o Al	t gg a Gl	a aa y As	c aa n Ly 57	S GI	a at y Il	t co e Pr	t ct o Le	eu G	ag gti lu Va 75	t 1728 I
•	gg G1 <sub>2</sub>	t aa y Ly	a gg s Gl	t ag y Ar 58	g at g Il 0	a tt e Le	g at u Il	t ga e Gl	g gg u G1 58	g ga y Gl 5	g ag u Ar	a gt g Va	g go	ct ct la Le 59	ca tr eu Lo 90	tg gga eu Gl	a 1776 y
	ta Ty	t gg r Gl	c to y Se 59	r Al	a gt a Va	g ca 1 Gl	g aa n As	c tg n Cy 60	's Le	g ga u As	t gc p Al	t gg a A	ıa ı	tt g le Va D5	tg c al L	ta ga eu Gl	a 1824 u
	tc Se	c cg r Ar 61	g G1	jc tt y Le	a ca eu Gl	a gt n Va	a ac il Th 61	ır Va	t go il Al	ia As	t go p Al eite	a Ai	gt t <sup>.</sup> rg Pl 20	tc t he C	gc a ys L	aa cc ys Pr	a 1872 o
					•					26	-116	04					

ctg Leu 625	gac Asp	cat His	gcc Ala	ctc Leu	ata Ile 630	agg Arg	agc Ser	ctt Leu	gca Ala	aaa Lys 635	tca Ser	cat His	gaa Glu	gtg Val	cta Leu 640	1920
	act Thr	gtc Val	gaa Glu	gaa Glu 645	gga Gly	tca Ser	att Ile	gga Gly	ggt Gly 650	ttt Phe	gga Gly	tct Ser	cat His	gtt Val 655	gtt Val	1968
cag Gln	ttc Phe	atg Met	gcc Ala 660	Leu	gat Asp	ggg Gly	ctt Leu	ctt Leu 665	gat Asp	ggc Gly	aag Lys	ttg Leu	aag Lys 670	tgg Trp	aga Arg	2016
cca Pro	ata Ile	gtt Val 675	Leu	cct Pro	gat Asp	cga Arg	tac Tyr 680	T 1 C	gac Asp	cat His	gga Gly	tct Ser 685	cct Pro	gtt Val	gat Asp	2064
caç Gli	ttg Leu 690	Ala	gaa Glu	gct Ala	ggc	cta Leu 695		cca Pro	tct Ser	cac His	att 11e 700	gca Ala	gca Ala	aca Thi	gta Val	2112
tt			a cti e Lei	gga u Gly	caa Glr 710	1 1111	aga Arg	gag Glu	gct Ala	cta Let 71:	a gaq u Gli	g gto u Val	ato Mei	g aca	a taa r	2160

<210> 58

<211> 719

<212> PRT

<213> Lycopersicon esculentum

·<400> 58

Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly 10 15

Val Val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile 20 25

is Gly Thr Asp Leu Gln Phe Leu Phe Gln His Lys Leu Thr His Glu 35 40 45

Val Lys Lys Arg Ser Arg Val Val Gln Ala Ser Leu Ser Glu Ser Gly 50 60

Glu Tyr Tyr Thr Gln Arg Pro Pro Thr Pro Ile Leu Asp Thr Val Asn 65 70 80

Tyr Pro Ile His Met Lys Asn Leu Ser Leu Lys Glu Leu Lys Gln Leu 85 90 95

Ala Asp Glu Leu Arg Ser Asp Thr Ile Phe Asn Val Ser Lys Thr Gly 100 100

Gly His Leu Gly Ser Ser Leu Gly Val Val Glu Leu Thr Val Ala Leu 125 Seite 85

His Tyr Val Phe Asn Ala Pro Gln Asp Arg Ile Leu Trp Asp Val Gly 130 135 140 His Gln Ser Tyr Pro His Lys Ile Leu Thr Gly Arg Arg Asp Lys Met 145 150 160 Ser Thr Leu Arg Gln Thr Asp Gly Leu Ala Gly Phe Thr Lys Arg Ser 165 170 175 Glu Ser Glu Tyr Asp Cys Phe Gly Thr Gly His Ser Ser Thr Thr Ile 180 185 190 Ser Ala Gly Leu Gly Met Ala Val Gly Arg Asp Leu Lys Gly Arg Asn 195 200 205 Asn Asn Val Ile Ala Val Ile Gly Asp Gly Ala Met Thr Ala Gly Gln 210 215 220 Ala Tyr Glu Ala Met Asn Asn Ala Gly Tyr Leu Asp Ser Asp Met Ile 225 230 240 Val Ile Leu Asn Asp Asn Arg Gln Val Ser Leu Pro Thr Ala Thr Leu 245 250 255 Asp Gly Pro Val Ala Pro Val Gly Ala Leu Ser Ser Ala Leu Ser Arg 260 265 270 Leu Gln Ser Asn Arg Pro Leu Arg Glu Leu Arg Glu Val Ala Lys Gly 275 280 285 Val Thr Lys Gln Ile Gly Gly Pro Met His Glu Leu Ala Ala Lys Val 290 295 300 sp Glu Tyr Ala Arg Gly Met Ile Ser Gly Ser Gly Ser Thr Leu Phe 5 . 310 320 Glu Glu Leu Gly Leu Tyr Tyr Ile Gly Pro Val Asp Gly His Asn Ile 325 330 335 Asp Asp Leu Ile Ala Ile Leu Lys Glu Val Arg Ser Thr Lys Thr Thr 340 345 . 350 Gly Pro Val Leu Ile His Val Val Thr Glu Lys Gly Arg Gly Tyr Pro 355 360 365 Tyr Ala Glu Arg Ala Ala Asp Lys Tyr His Gly Val Ala Lys Phe Asp 370 . 375 Pro Ala Thr Gly Lys Gln Phe Lys Ala Ser Ala Lys Thr Gln Ser Tyr 385 390 395 Seite 86

Thr Thr Tyr Phe Ala Glu Ala Leu Ile Ala Glu Ala Glu Ala Asp Lys 405. 410 415 Asp Ile Val Ala Ile His Ala Ala Met Gly Gly Gly Thr Gly Met Asn 420 425 430 Leu Phe His Arg Arg Phe Pro Thr Arg Cys Phe Asp Val Gly Ile Ala 435 440 445 Glu Gln His Ala Val Thr Phe Ala Ala Gly Leu Ala Cys Glu Gly Ile 450 455 460 Lys Pro Phe Cys Ala Ile Tyr Ser Ser Phe Met Gln Arg Ala Tyr Asp 465 470 480 Gln Val Val His Asp Val Asp Leu Gln Lys Leu Pro Val Arg Phe Ala 485 490 495 Met Asp Arg Ala Gly Leu Val Gly Ala Asp Gly Pro Thr His Cys Gly 500 505 510 Ala Phe Asp Val Thr Tyr Met Ala Cys Leu Pro Asn Met Val Val Met 515 525 Ala Pro Ser Asp Glu Ala Glu Leu Phe His Met Val Ala Thr Ala Ala 530 535 540 Ala Ile Asp Asp Arg Pro Ser Cys Phe Arg Tyr Pro Arg Gly Asn Gly 545 550 555 560 Ile Gly Val Glu Leu Pro Ala Gly Asn Lys Gly Ile Pro Leu Glu Val 565 570 575 ly Lys Gly Arg Ile Leu Ile Glu Gly Glu Arg Val Ala Leu Leu Gly 580 585 590 Tyr Gly Ser Ala Val Gln Asn Cys Leu Asp Ala Ala Ile Val Leu Glu 595 600 605 Ser Arg Gly Leu Gln Val Thr Val Ala Asp Ala Arg Phe Cys Lys Pro 610 615 620 Leu Asp His Ala Leu Ile Arg Ser Leu Ala Lys Ser His Glu Val Leu 625 630 635 640 Ile Thr Val Glu Glu Gly Ser Ile Gly Gly Phe Gly Ser His Val Val 645 650 655 Gln Phe Met Ala Leu Asp Gly Leu Leu Asp Gly Lys Leu Lys Trp Arg 660 665 670 Seite 87

Pro Ile Val Leu Pro Asp Arg Tyr Ile Asp His Gly Ser Pro Val Asp Gln Leu Ala Glu Ala Gly Leu Thr Pro Ser His Ile Ala Ala Thr Val Phe Asn Ile Leu Gly Gln Thr Arg Glu Ala Leu Glu Val Met Thr 705 710 715 <210> 59 <211> 1434 <212> DNA <213> Arabidopsis thaliana 20> <221> **CDS** <222> (1)..(1434)<223> <400> 59 atg atg aca tta aac tca cta tct cca gct gaa tcc aaa gct att tct Met Met Thr Leu Asn Ser Leu Ser Pro Ala Glu Ser Lys Ala Ile Ser 48 ttc ttg gat acc tcc agg ttc aat cca atc cct aaa ctc tca ggt ggg Phe Leu Asp Thr Ser Arg Phe Asn Pro Ile Pro Lys Leu Ser Gly Gly 20 25 30 96 ttt agt ttg agg agg agt caa ggg aga ggt ttt gga aaa ggt gtt Phe Ser Leu Arg Arg Arg Asn Gln Gly Arg Gly Phe Gly Lys Gly Val 35 40 45 144 g tgt tca gtg aaa gtg cag cag caa caa caa cct cct cca gca tgg ys Cys Ser Val Lys Val Gln Gln Gln Gln Pro Pro Pro Ala Trp 50 55 60 192 cct ggg aga gct gtc cct gag gcg cct cgt caa tct tgg gat gga cca Pro Gly Arg Ala Val Pro Glu Ala Pro Arg Gln Ser Trp Asp Gly Pro 65 70 75 80 240 aaa ccc atc tct atc gtt gga tct act ggt tct att ggc act cag aca Lys Pro Ile Ser Ile Val Gly Ser Thr Gly Ser Ile Gly Thr Gln Thr 85 90 95 288 ttg gat att gtg gct gag aat cct gac aaa ttc aga gtt gtg gct cta Leu Asp Ile Val Ala Glu Asn Pro Asp Lys Phe Arg Val Val Ala Leu 100 105 110 336 gct gct ggt tcg aat gtt act cta ctt gct gat cag gta agg aga ttt Ala Ala Gly Ser Asn Val Thr Leu Leu Ala Asp Gln Val Arg Arg Phe 115 120 125 384 125

aag cct gca ttg gtt gct gtt aga aac gag tca ctg att aat gag ctt

Seite 88

432

Lys	Pro 130	Ala	Leu	val	Аlа	Val 135	Arg		04seq Glu			Ile	Asn	Glu	Leu	
aaa Lys 145	gag Glu	gct Ala	tta Leu	gct Ala	gat Asp 150	ttg Leu	gac Asp	tat Tyr	aaa Lys	ctc Leu 155	gag Glu	att Ile	att Ile	cca Pro	gga Gly 160	480
gag Glu	caa Gln	gga Gly	gtg Val	att Ile 165	gag Glu	gtt Val	gcc Ala	cga Arg	cat His 170	cct Pro	gaa Glu	gct Ala	gta Val	acc Thr 175	gtt Val	528
gtt Val	acc Thr	gga Gly	ata Ile 180	gta Val	ggt Gly	tgt Cys	gcg Ala	gga Gly 185	cta Leu	aag Lys	cct Pro	acg Thr	gtt Val 190	gct Ala	gca Ala	576
att Ile	gaa Glu	gca Ala 195	gga Gly	aag Lys	gac Asp	att Ile	gct Ala 200	ctt Leu	gca Ala	aac Asn	aaa Lys	gag Glu 205	aca Thr	tta Leu	atc Ile	624
gca Ala	ggt Gly 210	ggt Gly	cct Pro	ttc Phe	gtg Val	ctt Leu 215	ccg Pro	ctt Leu	gcc Ala	aac Asn	aaa Lys 220	cat His	aat Asn	gta Val	aag Lys	672
t e 25	ctt Leu	ccg Pro	gca Ala	gat Asp	tca Ser 230	gaa Glu	cat His	tct Ser	gcc Ala	ata Ile 235	ttt Phe	cag Gln	tgt Cys	att Ile	caa Gln 240	720
ggt Gly	ttg Leu	cct Pro	gaa Glu	ggc Gly 245	gct Ala	ctg Leu	cgc Arg	aag Lys	ata Ile 250	atc Ile	ttg Leu	act Thr	gca Ala	tct Ser 255	ggt Gly	768
gga Gly	gct Ala	ttt Phe	agg Arg 260	gat Asp	tgg Trp	cct Pro	gtc Val	gaa Glu 265	aag Lys	cta Leu	aag Lys	gaa Glu	gtt Val 270	aaa Lys	gta Val	816
gcg Ala	gat Asp	gcg Ala 275	ttg Leu	aag Lys	cat His	cca Pro	aac Asn 280	tgg Trp	aac Asn	atg Met	gga Gly	aag Lys 285	aaa Lys	atc Ile	act Thr	864
gtg Val	gac Asp 290	Ser	gct Ala	acg Thr	ctt Leu	ttc Phe 295	Asn	aag Lys	ggt Gly	ctt Leu	gag Glu 300	gtc Val	att Ile	gaa Glu	gcg Ala	912
cat His 305	Tyr	ttg Leu	ttt Phe	gga Gly	gct Ala 310	Glu	tat Tyr	gac Asp	gat Asp	ata Ile 315	Glu	att Ile	gtc Val	att Ile	cat His 320	960
ro	caa Gln	agt Ser	atc Ile	ata Ile 325	cat His	tcc Ser	atg Met	att Ile	gaa Glu 330	aca Thr	cag Gln	gat Asp	tca Ser	tct Ser 335	Val	1008
ctt Leu	gct Ala	caa Gln	ttg Leu 340	Gly	tgg Trp	cct Pro	gat Asp	atg Met 345	Arg	tta Leu	ccg Pro	att Ile	ctc Leu 350	Tyr	acc Thr	1056
			Pro					Cys					Trp		aga Arg	1104
		Leu					' Ser					Lys			aat Asn	1152
gtg Val 385	Lys	tac Tyr	cca Pro	tco Ser	atg Met 390	: Asp	ctt Leu	gct Ala	tat Tyr	gct Ala 395	a Ala	gga Gly	cga Arg	gct Ala	gga Gly 400	1200
ggc	: aca	atg	act	: gga	gtt	cto	ago	gcc		aat te 8		ı aaa	gct	gtt	gaa	1248

<210> 60

<211> 477

12> PRT

<213> Arabidopsis thaliana

<400> 60

Met Met Thr Leu Asn Ser Leu Ser Pro Ala Glu Ser Lys Ala Ile Ser 1 10 15

Phe Leu Asp Thr Ser Arg Phe Asn Pro Ile Pro Lys Leu Ser Gly Gly 20 25 30

Phe Ser Leu Arg Arg Arg Asn Gln Gly Arg Gly Phe Gly Lys Gly Val 35 40 45

Lys Cys Ser Val Lys Val Gln Gln Gln Gln Pro Pro Pro Ala Trp 50 55 60

o Gly Arg Ala Val Pro Glu Ala Pro Arg Gln Ser Trp Asp Gly Pro 70 75 80

Lys Pro Ile Ser Ile Val Gly Ser Thr Gly Ser Ile Gly Thr Gln Thr  $85 \hspace{1cm} 90 \hspace{1cm} 95$ 

Leu Asp Ile Val Ala Glu Asn Pro Asp Lys Phe Arg Val Val Ala Leu 100 105 110

Ala Ala Gly Ser Asn Val Thr Leu Leu Ala Asp Gln Val Arg Arg Phe 115 120 125

Lys Pro Ala Leu Val Ala Val Arg Asn Glu Ser Leu Ile Asn Glu Leu 130 140

O4sequ.txt
Lys Glu Ala Leu Ala Asp Leu Asp Tyr Lys Leu Glu Ile Ile Pro Gly
145 150 155 160 Glu Gln Gly Val Ile Glu Val Ala Arg His Pro Glu Ala Val Thr Val 165 170 175 Val Thr Gly Ile Val Gly Cys Ala Gly Leu Lys Pro Thr Val Ala Ala 180 185 190 Ile Glu Ala Gly Lys Asp Ile Ala Leu Ala Asn Lys Glu Thr Leu Ile 195 200 205 Ala Gly Gly Pro Phe Val Leu Pro Leu Ala Asn Lys His Asn Val Lys 210 215 220 Ile Leu Pro Ala Asp Ser Glu His Ser Ala Ile Phe Gln Cys Ile Gln 225 230 235 240 Leu Pro Glu Gly Ala Leu Arg Lys Ile Île Leu Thr Ala Ser Gly 255 255 Gly Ala Phe Arg Asp Trp Pro Val Glu Lys Leu Lys Glu Val Lys Val 260 265 270 Ala Asp Ala Leu Lys His Pro Asn Trp Asn Met Gly Lys Lys Ile Thr 275 280 285 Val Asp Ser Ala Thr Leu Phe Asn Lys Gly Leu Glu Val Ile Glu Ala 290 295 300 His Tyr Leu Phe Gly Ala Glu Tyr Asp Asp Ile Glu Ile Val Ile His 305 . 310 315 320 Pro Gln Ser Ile Ile His Ser Met Ile Glu Thr Gln Asp Ser Ser Val 325 330 335 u Ala Gln Leu Gly Trp Pro Asp Met Arg Leu Pro Ile Leu Tyr Thr 340 345 350 Met Ser Trp Pro Asp Arg Val Pro Cys Ser Glu Val Thr Trp Pro Arg 355 360 365 Leu Asp Leu Cys Lys Leu Gly Ser Leu Thr Phe Lys Lys Pro Asp Asn 370 375 380 Val Lys Tyr Pro Ser Met Asp Leu Ala Tyr Ala Ala Gly Arg Ala Gly 385 390 395 Gly Thr Met Thr Gly Val Leu Ser Ala Ala Asn Glu Lys Ala Val Glu 405 410 415 Met Phe Ile Asp Glu Lys Ile Ser Tyr Leu Asp Ile Phe Lys Val Val 420 425 430 Glu Leu Thr Cys Asp Lys His Arg Asn Glu Leu Val Thr Ser Pro Ser 445 445 Leu Glu Glu Ile Val His Tyr Asp Leu Trp Ala Arg Glu Tyr Ala Ala 450 455 460 Asn Val Gln Leu Ser Ser Gly Ala Arg Pro Val His Ala 465 470 475 <210> 61 884 <211> <212> DNA <213> Adonis palaestina clone ApIPI28 <220> <221> CDS (180)..(884)<222> <223> <400> 61 cgtcgatcag gattaatcct ttatatagta tcttctccac caccactaaa acattatcag 60 cttcgtgttc ttctcccgct gttcatcttc agcagcgttg tcgtactctt tctatttctt 120 cttccatcac taacagtcct cgccgagggt tgaatcggct gttcgcctca acgtcgact 179 atg ggt gaa gtc gct gat gct ggt atg gat gcc gtc cag aag cgg ctt Met Gly Glu Val Ala Asp Ala Gly Met Asp Ala Val Gln Lys Arg Leu 1 5 10 227 eg ttc gac gat gaa tgt att ttg gtg gat gag aat gac aag gtc gtc het Phe Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val 20 25 30 275 gga cat gat tcc aaa tac aac tgt cat ttg atg gaa aag ata gag gca Gly His Asp Ser Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala 35 40 45323 gaa aac ttg ctt cac aga gcc ttc agt gtt ttc tta ttc aac tca aaa Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys 50 55 60 371 tac gag ttg ctt ctt cag caa cga tct gca acg aag gta aca ttc ccg Tyr Glu Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro 65 70 75 80 419 ctc gta tgg aca aac acc tgt tgc agc cat ccc ctc ttc cgt gat tcc Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Phe Arg Asp Ser 85 90 95 467

gaa ctc ata gaa gaa aat ttt ctc ggg gta cga aac gct gca caa agg

Seite 92

515

									·	14sec	ļu.τ>	(T	_		_		
	Glu	Leu	Ile	Glu 100	Glu	Asn	Phe	Leu	Gly 105	Val	Arg	Asn	Ala	Ala 110	Gln	Arg	
	aag Lys	ctt Leu	tta Leu 115	gac Asp	gag Glu	cta Leu	ggc Gly	att Ile 120	cca Pro	gct Ala	gaa Glu	gac Asp	gta Val 125	cca Pro	gtt Val	gat Asp	563
	gaa Glu	ttc Phe 130	act Thr	cct Pro	ctt Leu	ggt Gly	cgc Arg 135	att Ile	ctt Leu	tac Tyr	aaa Lys	gct Ala 140	cca Pro	tct Ser	gac Asp	gga Gly	611
	aaa Lys 145	tgg Trp	gga Gly	gag Glu	cac His	gaa Glu 150	ctg Leu	gac Asp	tat Tyr	ctt Leu	ctg Leu 155	ttt Phe	att Ile	gtc Val	cga Arg	gat Asp 160	659
	gtg Val	aaa Lys	tac Tyr	gat Asp	cca Pro 165	aac Asn	cca Pro	gat Asp	gaa Glu	gtt Val 170	gct Ala	gac Asp	gct Ala	aag Lys	tac Tyr 175	gtt Val	707
	aat Asn	cgc Arg	gag Glu	gag Glu 180	ttg Leu	aaa Lys	gag Glu	ata Ile	ctg Leu 185	aga Arg	aaa Lys	gct Ala	gat Asp	gca Ala 190	GIY	gaa Glu	755
	g	gga Gly	ata Ile 195	Lys	ttg Leu	tct Ser	cct Pro	tgg Trp 200	Phe	aga Arg	ttg Leu	gtt Val	gtg Val 205	ASP	aac Asn	ttt Phe	803
•	ttg Leu	ttc Phe 210	Lys	tgg Trp	tgg Trp	gat Asp	cat His 215	Val	gag Glu	gag Glu	ggg Gly	aag Lys 220	lle	aag Lys	gac Asp	gtc Val	851
	gcc Ala 225	Äsp	ato Met	g aaa Lys	act Thr	ato Ile 230	: His	aag Lys	ttg Leu	act Thr	taa	ι					884
	<21	.0>	62														
	<21	.1>	234														
		L2>															
	<21	L3>	Ado	nis į	palae	estir	na c̄	lone	.ApII	PI28							
_		20.	<b>C</b> 2														

Asp Glu Val Ala Asp Ala Gly Met Asp Ala Val Gln Lys Arg Leu Met Phe Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val Gly His Asp Ser Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Glu Leu Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro 80

O4sequ.txt Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Phe Arg Asp Ser 85 90 95	
Glu Leu Ile Glu Glu Asn Phe Leu Gly Val Arg Asn Ala Ala Gln Arg 100 105 110	
Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp 115 120 125	
Glu Phe Thr Pro Leu Gly Arg Ile Leu Tyr Lys Ala Pro Ser Asp Gly 130 135 140	
Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Ile Val Arg Asp 145 150 155 160	
Val Lys Tyr Asp Pro Asn Pro Asp Glu Val Ala Asp Ala Lys Tyr Val 165 170 175	
n Arg Glu Glu Leu Lys Glu Ile Leu Arg Lys Ala Asp Ala Gly Glu 180 185 190	
Glu Gly Ile Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe 195 200 205	
Leu Phe Lys Trp Trp Asp His Val Glu Glu Gly Lys Ile Lys Asp Val 210 215 220	
Ala Asp Met Lys Thr Ile His Lys Leu Thr 225 230	
<210> 63	
<211> 1402	
<212> DNA	
213> Arabidopsis thaliana	
<220>	
<221> CDS	
<222> (52)(1317)	
<223>	
<400> 63 aagtctttgc ctctttggtt tactttcctc tgttttcgat ccatttagaa a atg tta 57 Met Leu 1	
ttc acg agg agt gtt gct cgg att tct tct aag ttt ctg aga aac cgt 105 Phe Thr Arg Ser Val Ala Arg Ile Ser Ser Lys Phe Leu Arg Asn Arg 5 10 15 Seite 94	

	agc Ser	ttc Phe 20	tat Tyr	ggc Gly	tcc Ser	ser	caa Gln 25	tct Ser	ctc Leu	gcc Ala	tct Ser	cat His 30	cgg Arg	ttc Phe	gca Ala	atc Ile	153
	att Ile 35	ccc Pro	gat Asp	cag Gln	ggt Gly	cac His 40	tct Ser	tgt Cys	tct Ser	gac Asp	tct Ser 45	cca Pro	cac His	aag Lys	ggt Gly	tac Tyr 50	201
	gtt Val	tgc Cys	aga Arg	aca Thr	act Thr 55	tat Tyr	tca Ser	ttg Leu	aaa Lys	tct ser 60	ccg Pro	gtt Val	ttt Phe	ggt Gly	gga Gly 65	ttt Phe	249
	agt Ser	cat His	caa Gln	ctc Leu 70	tat Tyr	cac His	cag Gln	agt Ser	agc Ser 75	tcc Ser	ttg Leu	gtt Val	gag Glu	gag Glu 80	gag Glu	ctt Leu	297
	gac Asp	cca Pro	ttt Phe 85	tcg Ser	ctt Leu	gtt Val	gcc Ala	gat Asp 90	gag Glu	ctg Leu	tca Ser	ctt Leu	ctt Leu 95	agt Ser	aat Asn	aag Lys	345
4	ttg Leu	aga Arg 100	gag Glu	atg Met	gta Val	ctt Leu	gcc Ala 105	gag Glu	gtt Val	cca Pro	aag Lys	ctt Leu 110	gcc Ala	tct Ser	gct Ala	gct Ala	393
	g Glu 115	Tyr	ttc Phe	ttc Phe	aaa Lys	agg Arg 120	ggt Gly	gtg Val	caa Gln	gga Gly	aaa Lys 125	cag Gln	ttt Phe	cgt Arg	tca Ser	act Thr 130	441
	att Ile	ttg Leu	ctg Leu	ctg Leu	atg Met 135	Ala	aca Thr	gct Ala	ctg Leu	gat Asp 140	Val	cga Arg	gtt Val	cca Pro	gaa Glu 145	gca Ala	489
	ttg Leu	att Ile	ggg Gly	gaa Glu 150	Ser	aca Thr	gat Asp	ata Ile	gtc Val 155	Thr	tca Ser	gaa Glu	tta Leu	cgc Arg 160	Val	agg Arg	537
	caa Gln	cgg Arg	ggt Gly 165	Ile	gct Ala	gaa Glu	atc	act Thr 170	Glu	atg Met	ata Ile	cac His	gtc Val 175	Ala	agt Ser	cta Leu	585
	ctg Lei	cac His 180	Asp	gat Asp	gto Val	ttg Leu	gat Asp 185	Asp	gcc Ala	gat Asp	aca Thr	agg Arg 190	g Arg	ggt Gly	gtt Va	ggt Gly	633
	tco	· Lei	ı aat ı Asr	gti Va	gta Val	atg Met 200	: Gly	aac Asr	aag 1 Lys	atg Met	tcg Ser 205	· Va	a tta l Leu	ı gca ı Ala	a gga a Gly	a gac / Asp 210	681
	tto Phe	ttg e Lei	j cto i Lei	tce Sei	c cgg r Arg 215	) Ala	tgt Cys	ggg G1y	g gct / Ala	cto Lei 220	ı Ala	gci a Ala	t tta a Leu	a aag u Lys	g aa s Ası 22	c aca n Thr 5	729
	gaç Gli	g gti u Va	t gta l Va	a gca l Ala 230	a Lei	a ctt u Lei	gca ı Ala	a act	t gc1 r Ala 23	a Va	a gaa l Glu	a ca u Hi	t cti s Lei	t gt: u Va 240	i Th	c ggt r Gly	777
	ga: Gl:	a ac u Th	r Ate	<u>t</u> Gl	g ata u Ile	a act e Thi	ag Se	t tca r se 250	r Th	c gag r Gli	g cag u Gli	g cg n Ar	t ta g Ty: 25	r se	t at r Me	g gac t Asp	825
	ta Ty	c ta r Ty 26	r Me	g ca t Gl	g aag n Ly:	g aca s Th	a ta r Ty 26	r Ty	t aa r Ly	g ac s Th	a gc r Al	a tc a se 27	r Le	a at u Il	c tc e Se	t aac r Asn	873
	ag Se 27	ŗ Cÿ	c aa s Ly	a gc s Al	t gt a Va	t gc 1 Al: 28	a Va	t ct 1 Le	c ac u Th	r Gl	y G1 28	n Th 5	a gc r Al	a ga a Gl	a gt u Va	t gcc 1 Ala 290	921
	•									56	ite	33					

gtg Val	tta Leu	gct Ala	ttt Phe	gag Glu 295	tat Tyr	ggg Gly	agg Arg	aat Asn	ctg Leu 300	ggt Gly	tta Leu	gca Ala	ttc Phe	caa Gln 305	tta Leu	969
ata Ile	gac Asp	gac Asp	att Ile 310	ctt Leu	gat Asp	ttc Phe	acg Thr	ggc Gly 315	aca Thr	tct Ser	gcc Ala	tct Ser	ctc Leu 320	gga Gly	aag Lys	1017
gga Gly	tcg Ser	ttg Leu 325	tca Ser	gat Asp	att Ile	cgc Arg	cat His 330	gga Gly	gtc Val	ata Ile	aca Thr	gcc Ala 335	cca Pro	atc Ile	ctc Leu	1065
ttt Phe	gcc Ala 340	atg Met	gaa Glu	gag Glu	ttt Phe	cct Pro 345	caa Gln	cta Leu	cgc Arg	gaa Glu	gtt Val 350	gtt Val	gat Asp	caa Gln	gtt Val	1113
gaa Glu 355	aaa Lys	gat Asp	cct Pro	agg Arg	aat Asn 360	gtt Val	gac Asp	att Ile	gct Ala	tta Leu 365	gag Glu	tat Tyr	ctt Leu	ggg Gly	aag Lys 370	1161
agc Ser	aag Lys	gga Gly	ata Ile	cag Gln 375	agg Arg	gca Ala	aga Arg	gaa Glu	tta Leu 380	gcc Ala	atg Met	gaa Glu	cat His	gcg Ala 385	aat Asn	1209
_ca Leu	gca Ala	gca Ala	gct Ala 390	gca Ala	atc Ile	ggg Gly	tct Ser	cta Leu 395	cct Pro	gaa Glu	aca Thr	gac Asp	aat Asn 400	Glu	gat Asp	1257
gtc Val	aaa Lys	aga Arg 405	ser	agg Arg	cgg Arg	gca Ala	ctt Leu 410	att Ile	gac Asp	ttg Leu	acc Thr	cat His 415	Arg	gtc Val	atc Ile	1305
		Asn	aag Lys		gatt	aag	taat	gttt	ct c	tcta	taca	.c ca	aaac	atto		1357
ctc	attt	cat	ttgt	agga	tt t	tgtt	ggtc	c aa	ttcg	tttc	acg	jaa				1402

<210> 64

<211> 422

<212> PRT

213> Arabidopsis thaliana

<400> 64

Met Leu Phe Thr Arg Ser Val Ala Arg Ile Ser Ser Lys Phe Leu Arg 10 15

Asn Arg Ser Phe Tyr Gly Ser Ser Gln Ser Leu Ala Ser His Arg Phe 20 25 30

Ala Ile Ile Pro Asp Gln Gly His Ser Cys Ser Asp Ser Pro His Lys 40 45

Gly Tyr Val Cys Arg Thr Thr Tyr Ser Leu Lys Ser Pro Val Phe Gly 50 60

Gly Phe Ser His Gln Leu Tyr His Gln Ser Ser Leu Val Glu Glu 65 70 75 80 Glu Leu Asp Pro Phe Ser Leu Val Ala Asp Glu Leu Ser Leu Leu Ser 85 90 95 Asn Lys Leu Arg Glu Met Val Leu Ala Glu Val Pro Lys Leu Ala Ser 100 105 110 Ala Ala Glu Tyr Phe Phe Lys Arg Gly Val Gln Gly Lys Gln Phe Arg 115 120 125 Ser Thr Ile Leu Leu Met Ala Thr Ala Leu Asp Val Arg Val Pro 130 135 140 Glu Ala Leu Ile Gly Glu Ser Thr Asp Ile Val Thr Ser Glu Leu Arg 145 150 155 160 l Arg Gln Arg Gly Ile Ala Glu Ile Thr Glu Met Ile His Val Ala 165 170 175 Ser Leu Leu His Asp Asp Val Leu Asp Asp Ala Asp Thr Arg Arg Gly 180 185 190 Val Gly Ser Leu Asn Val Val Met Gly Asn Lys Met Ser Val Leu Ala 195 200 205 Gly Asp Phe Leu Leu Ser Arg Ala Cys Gly Ala Leu Ala Ala Leu Lys 210 215 220 Asn Thr Glu Val Val Ala Leu Leu Ala Thr Ala Val Glu His Leu Val 225 230 235 240 Thr Gly Glu Thr Met Glu Ile Thr Ser Ser Thr Glu Gln Arg Tyr Ser 245 250 255 et Asp Tyr Tyr Met Gln Lys Thr Tyr Tyr Lys Thr Ala Ser Leu Ile 260 265 270 Ser Asn Ser Cys Lys Ala Val Ala Val Leu Thr Gly Gln Thr Ala Glu 275 280 285 Val Ala Val Leu Ala Phe Glu Tyr Gly Arg Asn Leu Gly Leu Ala Phe 290 295 300 Gln Leu Ile Asp Asp Ile Leu Asp Phe Thr Gly Thr Ser Ala Ser Leu 305 310 315 Gly Lys Gly Ser Leu Ser Asp Ile Arg His Gly Val Ile Thr Ala Pro 325 330 335

04sequ.txt Ile Leu Phe Ala Met Glu Glu Phe Pro Gln Leu Arg Glu Val Val Asp Gln Val Glu Lys Asp Pro Arg Asn Val Asp Ile Ala Leu Glu Tyr Leu 355 360 365 Gly Lys Ser Lys Gly Ile Gln Arg Ala Arg Glu Leu Ala Met Glu His 370 375 380 Ala Asn Leu Ala Ala Ala Ile Gly Ser Leu Pro Glu Thr Asp Asn 385 390 395 Glu Asp Val Lys Arg Ser Arg Arg Ala Leu Ile Asp Leu Thr His Arg 405 410 415 Val Ile Thr Arg Asn Lys 420 10> 65 <211> 1155 <212> DNA <213> Arabidopsis thaliana <220> <221> **CDS** (1)..(1155)<222> <223> <400> eatg agt gtg agt tgt tgt tgt agg aat ctg ggc aag aca ata aaa aag et Ser Val Ser Cys Cys Cys Arg Asn Leu Gly Lys Thr Ile Lys Lys 5 10 15 gca ata cct tca cat cat ttg cat ctg aga agt ctt ggt ggg agt ctc Ala Ile Pro Ser His His Leu His Leu Arg Ser Leu Gly Gly Ser Leu 20 25 . 30 tat cgt cgt cgt atc caa agc tct tca atg gag acc gat ctc aag tca Tyr Arg Arg Arg Ile Gln Ser Ser Met Glu Thr Asp Leu Lys Ser 35 40 45

tat cgt cgt atc caa agc tct tca atg gag acc gat ctc aag tca Ser Met Glu Thr Asp Leu Lys Ser 40

acc ttt ctc aac gtt tat tct gtt ctc aag tct gac ctt ctt cat gac Leu Lys Ser Asp Soo Ser Phe Glu Phe Thr Asp Glu Ser Arg Leu Trp Val Asp Arg Met Roll Ser Val Leu Asp Tyr Asp Ctc aat cgt gac ctc tct tct gat gat cgg atg Ctc aat cgg ggt ctc tct gg gtt gat cgg atg Asp Arg Met Roll Ser Arg Leu Asp Arg Gly Gly Lys Leu Asp Arg Gly Leu Ser Val Ser Val Ser Val Ser Ser Leu Asp Arg Gly Gly Lys Leu Asp Arg Gly Leu Ser Val S

48

96

gtt Val	gac Asp	agt Ser	ttc Phe 100	aaa Lys	ctt Leu	ttg Leu	aag Lys	caa Gln 105	ggc Gly	aat Asn	gat Asp	Leu	act Thr 110	gag Glu	caa Gln	336
gag Glu	gtt Val	ttc Phe 115	ctc Leu	tct Ser	tgt Cys	gct Ala	ctc Leu 120	ggt Gly	tgg Trp	tgc Cys	att Ile	gaa Glu 125	tgg Trp	ctc Leu	caa Gln	384
gct Ala	tat Tyr 130	ttc Phe	ctt Leu	gtg Val	ctt Leu	gat Asp 135	gat Asp	att Ile	atg Met	gat Asp	aac Asn 140	tct Ser	gtc Val	act Thr	cgc Arg	432
cgt Arg 145	ggt Gly	caa Gln	cct Pro	tgc Cys	tgg Trp 150	ttc Phe	aga Arg	gtt Val	cct Pro	cag Gln 155	gtt Val	ggt Gly	atg Met	gtt Val	gcc Ala 160	480
atc Ile	aat Asn	gat Asp	ggg Gly	att Ile 165	cta Leu	ctt Leu	cgc Arg	aat Asn	cac His 170	atc Ile	cac His	agg Arg	att Ile	ctc Leu 175	aaa Lys	528
aag Lys	cat His	ttc Phe	cgt Arg 180	gat Asp	aag Lys	cct Pro	tac Tyr	tat Tyr 185	gtt Val	gac Asp	ctt Leu	gtt Val	gat Asp 190	ttg Leu	ttt Phe	576
at Asn	gag Glu	gtt Val 195	gag Glu	ttg Leu	caa Gln	aca Thr	gct Ala 200	tgt Cys	ggc Gly	cag Gln	atg Met	ata Ile 205	gat Asp	ttg Leu	atc Ịle	624
acc Thr	acc Thr 210	Phe	gaa Glu	gga Gly	gaa Glu	aag Lys 215	gat Asp	ttg Leu	gcc Ala	aag Lys	tac Tyr 220	tca Ser	ttg Leu	tca Ser	atc Ile	672
cac His 225	Arg	cgt Arg	att Ile	gtc Val	cag Gln 230	tac Tyr	aaa Lys	acg Thr	gct Ala	tat Tyr 235	Tyr	tca Ser	ttt Phe	tat Tyr	ctc Leu 240	720
cct Pro	gtt Val	gct Ala	tgt Cys	gcg Ala 245		ctt Leu	atg Met	gcg Ala	ggc Gly 250	gaa Glu	aat I Asn	ttg Leu	gaa Glu	aac Asn 255	cat His	768
att Ile	gac Asp	gtg Val	aaa Lys 260	Asn	gtt Val	ctt Leu	gtt Val	gac Asp 265	Met	gga Gly	a atc / Ile	tac Tyr	ttc Phe 270	Glr	gtg Val	816
cag In	gat Asp	gat Asp 275	Tyr	ctg Leu	gat Asp	tgt Cys	ttt Phe 280	: Āla	gat Asp	cco Pro	gag Gli	acg Thr 285	: Lei	ggo Gly	aag / Lys	864
ata Ile	gga Gly 290	/ Thr	gat Asp	ata Ile	gaa Glu	gat Asp 295	Phe	aaa Lys	tgo Cys	tce Sei	g tgg r Trp 300	Lei	g gto i Val	gti Va	t aag l Lys	912
gca Ala 305	Lei	a gag u Gli	ı Arg	tgc g Cys	ago Ser 310	'G]ι	ı gaa ı Glu	r caa I Glr	a act	t aag r Ly: 31	s Ile	a tta e Lei	a tai u Tyi	t gae r Gli	g aac u Asn 320	960
tat Tyr	gg1	t aaa y Ly:	a cco s Pro	c gad o Asp 325	Pro	tco Sei	g aad Ast	gti Va	gc1 1 A1a 330	a Ly	a gtç s Va	aaq Ly:	g ga <sup>.</sup> s As <sub>l</sub>	t cte p Le 33	c tac u Tyr 5	1008
aaa Lys	a ga s Gli	g cte	g ga u Ası 34	p Lei	t gag u Gli	g gga u Gly	a gti y Va	t tte l Phe 34	e Me	g ga t Gl	g ta u Ty	t gag r Gli	g ag u se 35	r Ly	a agc s Ser	1056
tac Tyi	ga r Gl	g aa u Ly: 35	s Le	g act u Th	t gga r Gly	a gce y Ala	g at a Il 36	e Gi	u GI	у Ні	s GI	a ag n Se 36	<u>r</u> Ly	a gc s Al	a atc a Ile	1104
									se	ite	99					

caa gca gtg cta aaa tcc ttc ttg gct aag atc tac aag agg cag aag
Gln Ala Val Leu Lys Ser Phe Leu Ala Lys Ile Tyr Lys Arg Gln Lys
370
375
380

1155

<210> 66

<211> 384

<212> PRT

<213> Arabidopsis thaliana

<400> 66

Met Ser Val Ser Cys Cys Cys Arg Asn Leu Gly Lys Thr Ile Lys Lys 1 10 15

l Ile Pro Ser His His Leu His Leu Arg Ser Leu Gly Gly Ser Leu 20 25 30

Tyr Arg Arg Arg Ile Gln Ser Ser Ser Met Glu Thr Asp Leu Lys Ser 35 40 45

Thr Phe Leu Asn Val Tyr Ser Val Leu Lys Ser Asp Leu Leu His Asp 50 60

Pro Ser Phe Glu Phe Thr Asn Glu Ser Arg Leu Trp Val Asp Arg Met 65 70 75 80

Leu Asp Tyr Asn Val Arg Gly Gly Lys Leu Asn Arg Gly Leu Ser Val 85 90 95

Val Asp Ser Phe Lys Leu Leu Lys Gln Gly Asn Asp Leu Thr Glu Gln
100 105 110

u Val Phe Leu Ser Cys Ala Leu Gly Trp Cys Ile Glu Trp Leu Gln 115 120 125

Ala Tyr Phe Leu Val Leu Asp Asp Ile Met Asp Asn Ser Val Thr Arg 130 135 140

Arg Gly Gln Pro Cys Trp Phe Arg Val Pro Gln Val Gly Met Val Ala 145 150 160

Ile Asn Asp Gly Ile Leu Leu Arg Asn His Ile His Arg Ile Leu Lys 165 170 175

Lys His Phe Arg Asp Lys Pro Tyr Tyr Val Asp Leu Val Asp Leu Phe 180 185 190 O4sequ.txt
Asn Glu Val Glu Leu Gln Thr Ala Cys Gly Gln Met Ile Asp Leu Ile
195 200 205 Thr Thr Phe Glu Gly Glu Lys Asp Leu Ala Lys Tyr Ser Leu Ser Ile 210 215 220 His Arg Arg Ile Val Gln Tyr Lys Thr Ala Tyr Tyr Ser Phe Tyr Leu 225 230 235 240 Pro Val Ala Cys Ala Leu Leu Met Ala Gly Glu Asn Leu Glu Asn His 245 250 255 Ile Asp Val Lys Asn Val Leu Val Asp Met Gly Ile Tyr Phe Gln Val 260 265 270 Gln Asp Asp Tyr Leu Asp Cys Phe Ala Asp Pro Glu Thr Leu Gly Lys 275 280 285 Gly Thr Asp Ile Glu Asp Phe Lys Cys Ser Trp Leu Val Val Lys 290 295 300 Ala Leu Glu Arg Cys Ser Glu Glu Gln Thr Lys Ile Leu Tyr Glu Asn 305 310 315 Leu Tyr Glu Asn 320Tyr Gly Lys Pro Asp Pro Ser Asn Val Ala Lys Val Lys Asp Leu Tyr 325 330 335 Lys Glu Leu Asp Leu Glu Gly Val Phe Met Glu Tyr Glu Ser Lys Ser 340 345 350 Tyr Glu Lys Leu Thr Gly Ala Ile Glu Gly His Gln Ser Lys Ala Ile 355 360 365 Gln Ala Val Leu Lys Ser Phe Leu Ala Lys Ile Tyr Lys Arg Gln Lys 370 375 380 10> 67 <211> 1101 <212> DNA

<213> Sinabs alba

<220>

<221> CDS

<222> (1)..(1101)

<223>

	-400	. 6	57						_			-			00> 67 get tet tea gtg act eet eta ggt tea tgg gtt ett ett eac eat 48													
	ata	act	tct	tca Ser	gtg Val 5	act Thr	cct Pro	cta Leu	Gly	tca Ser 10	tgg Trp	gtt Val	ctt Leu	ctt Leu	cac His 15	cat His	48											
	cat His	cct Pro	tca Ser	act Thr 20	atc Ile	tta Leu	acc Thr	caa Gln	tcc ser 25	aga Arg	tcc Ser	aga Arg	tct Ser	cct Pro 30	cct Pro	tct Ser	96											
	ctc Leu	atc Ile	acc Thr 35	ctt Leu	aaa Lys	ccc Pro	atc Ile	tcc Ser 40	ctc Leu	act Thr	cca Pro	aaa Lys	cgc Arg 45	acc Thr	gtt Val	tcg Ser	144											
	tct Ser	tct Ser 50	tcc Ser	tcc Ser	tct Ser	tcc Ser	ctc Leu 55	atc Ile	acc Thr	aaa Lys	gaa Glu	gac Asp 60	aac Asn	aac Asn	ctc Leu	aaa Lys	192											
	tcc Ser 65	tct Ser	tcc Ser	tct Ser	tcc Ser	ttc Phe 70	gat Asp	ttc Phe	atg Met	tct Ser	tac Tyr 75	atc Ile	atc Ile	cgc Arg	aaa Lys	gcc Ala 80	240											
•	gac Asp	tcc Ser	gtc Val	aac Asn	aaa Lys 85	gcc Ala	tta Leu	gac Asp	tcc Ser	gcc Ala 90	gtc Val	cct Pro	ctc Leu	cgg Arg	gag Glu 95	cca Pro	288											
	Leu	aag Lys	atc Ile	cac His 100	Glu	gcg Ala	atg Met	cgt Arg	tac Tyr 105	tct Ser	ctc Leu	ctc Leu	gcc Ala	gga Gly 110	Gly	aaa Lys	336											
	cgc Arg	gtc val	aga Arg 115	cca Pro	gtt Val	ctc Leu	tgc Cys	atc Ile 120	gcc Ala	gcg Ala	tgc Cys	gag Glu	cta Leu 125	٧a٦	gga Gly	gga Gly	384											
	gaa Glu	gag Glu 130	ı Ser	tta Leu	gct Ala	atg Met	ccg Pro 135	Ala	cgt Arg	tgc Cys	gcc Ala	gtg Val 140	Glu	atg Met	ato :Ile	cac His	432											
	acc Thr 145	Met	tcg Ser	ttg Leu	ato Ile	cac His 150	Asp	gac Asp	ttg Leu	cct Pro	tgt Cys 155	Met	gat Asp	aac Asr	gac S Asp	gat Asp 160	480											
	ctc Leu	cgo	c cgc g Arg	gga Gly	aag Lys 165	Pro	acg Thr	aat Asn	cac His	aaa Lys 170	· val	tac Tyr	ggc Gly	gaa Glu	gad Asp 175	gtg Val	528											
	gcg	gti Va	t tta i Leu	gco Ala 180	ı Gly	gac / Asp	gcg Ala	ctt Leu	ctt Leu 185	Ser	tto Phe	gco	tto Phe	gaç Gli 190	i His	t tta s Leu	576											
	gcg	tco Sei	g gct r Ala 195	1 Thi	g ago r Sei	tcg Ser	gag Gli	g gtt i val 200	Ser	ccg Pro	g gcg Ala	g aga a Arg	g gtg g Va 20!	<u>l</u> va	t aga l Ar	a gct g Ala	624											
	gto Val	gg   Gl   21	y Glu	g ttg u Lei	g gct u Ala	t aaa a Lys	g gco s Ala 21	a Ile	c ggc e Gly	aco Thi	gaa Glu	a ggg u Gly 220	y Lei	c gtg u Va	g gc	g gga a Gly	672											
	caa Glr 225	ı Va	g gtg I Va	g ga I Ası	t ata p Ilo	a ago e Seo 230	r Sei	t gaa r Glu	a ggg u Gly	tto Lei	g gae u As <sub>l</sub> 23	p Lei	a aa u As	c aa n As	c gt n Va	c gga 1 Gly 240	720											
	tt <u>g</u> Lei	g ga u Gl	g ca u Hi:	t tt	g aa u Ly 24	s Ph	t ata e Il	a cat e Hi:	t ttg s Lei	g ca u Hi: 250	S Ly	a ac s Th	g gc r Al	g gc a Al	g tt a Le 25	g ctt u Leu 5	768											
	ga: Gl:	a gc u Al	t tc a Se	a gc r Al 26	g gt a Va 0	t tte	g gg u Gl	t gg y Gl	g ato y Ile 26	5			a gg y Gl	g ag y se 27	t ga r As 0	t gaa p.Glu	816											
										Set	ite :	102																

ga G1	ig a	tc []e	gag Glu 275	agg Arg	ctg Leu	agg Arg	aag Lys	ttc Phe 280	gcg Ala	agg Arg	tgt Cys	att Ile	ggg Gly 285	ttg Leu	ttg Leu		<b>86</b> 4	ŀ
ca G1	n v	gtg /al 290	gtt val	gat Asp	gat Asp	atc Ile	ttg Leu 295	gac Asp	gtg Val	acg Thr	aaa Lys	tcg Ser 300	tct Ser	caa Gln	gaa Glu	ctg Leu	912	2
gç G1 30	V L	aaa .ys	acc Thr	gct Ala	ggg Gly	aaa Lys 310	gat Asp	ttg Leu	att Ile	gct Ala	gat Asp 315	aag Lys	ttg Leu	act Thr	tat Tyr	ccg Pro 320	960	)
a a Ly	ig c	tc eu	atg Met	ggt Gly	ttg Leu 325	gag Glu	aaa Lys	tcg Ser	aga Arg	gag Glu 330	ttc Phe	gct Ala	gag Glu	aag Lys	ttg Leu 335	aat Asn	1008	3
ac Tł	a g ir G	gag Slu	gca Ala	cgt Arg 340	gat Asp	cag Gln	ctt Leu	tta Leu	ggg Gly 345	ttt Phe	gat Asp	tcc Ser	gac Asp		gtt Val		1056	5
CC Pr	ct t	ttg _eu	ttg Leu 355	gct Ala	ttg Leu	gct Ala	aat Asn	tac Tyr 360	att Ile	gcc Ala	aat Asn	aga Arg	cag Gln 365	aac Asn	tga		110	L

<210> 68

<211> 366

<212> PRT

<213> Sinabs alba

<400> 68

Met Ala Ser Ser Val Thr Pro Leu Gly Ser Trp Val Leu Leu His His 10 15

His Pro Ser Thr Ile Leu Thr Gln Ser Arg Ser Arg Ser Pro Pro Ser 20 25 30

eu Ile Thr Leu Lys Pro Ile Ser Leu Thr Pro Lys Arg Thr Val Ser 35 40 45

Ser Ser Ser Ser Ser Leu Ile Thr Lys Glu Asp Asn Asn Leu Lys 50 60

Ser Ser Ser Ser Phe Asp Phe Met Ser Tyr Ile Ile Arg Lys Ala 65 70 75 80

Asp Ser Val Asn Lys Ala Leu Asp Ser Ala Val Pro Leu Arg Glu Pro 85 90 95

Leu Lys Ile His Glu Ala Met Arg Tyr Ser Leu Leu Ala Gly Gly Lys 100 105 110

Arg Val Arg Pro Val Leu Cys Ile Ala Ala Cys Glu Leu Val Gly Gly
115 . Seite 103

Glu Glu Ser Leu Ala Met Pro Ala Arg Cys Ala Val Glu Met Ile His

Thr Met Ser Leu Ile His Asp Asp Leu Pro Cys Met Asp Asn Asp Asp 160

Leu Arg Arg Gly Lys Pro Thr Asn His Lys Val Tyr Gly Glu Asp Val 175

Ala Val Leu Ala Gly Asp Ala Leu Leu Ser Phe Ala Phe Glu His Leu 180

Ala Ser Ala Thr Ser Ser Glu Val Ser Pro Ala Arg Val Val Arg Ala Val Gly Glu Ceu Ala Gly 210

Gln Val Val Asp Ile Ser Ser Glu Gly Leu Asp Leu Asn Asn Val Gly 240

Leu Glu His Leu Asp Ile Ser Ser Glu Gly Leu Asp Leu Asn Asn Val Gly 240

Leu Glu His Leu Lys Phe Ile His Leu His Lys Thr Ala Ala Leu Leu 245 250 255

Glu Ala Ser Ala Val Leu Gly Gly Ile Ile Gly Gly Gly Ser Asp Glu 260 265 270

Glu Ile Glu Arg Leu Arg Lys Phe Ala Arg Cys Ile Gly Leu Leu Phe 275 280 285

Gln Val Val Asp Asp Ile Leu Asp Val Thr Lys Ser Ser Gln Glu Leu 290 295 300

ly Lys Thr Ala Gly Lys Asp Leu Ile Ala Asp Lys Leu Thr Tyr Pro 310 315 320

Lys Leu Met Gly Leu Glu Lys Ser Arg Glu Phe Ala Glu Lys Leu Asn 325 330 335

Thr Glu Ala Arg Asp Gln Leu Leu Gly Phe Asp Ser Asp Lys Val Ala 340 350

Pro Leu Leu Ala Leu Ala Asn Tyr Ile Ala Asn Arg Gln Asn 355 360 365

<210> 69

<211> 930

<212> DNA

<213> Erwinia uredovora <220> <221> **CDS** <222> (1)..(930)<223> <400> 69

atg aat aat ccg tcg tta ctc aat cat gcg gtc gaa acg atg gca gtt Met Asn Asn Pro Ser Leu Leu Asn His Ala Val Glu Thr Met Ala Val 48 ggc tcg aaa agt ttt gcg aca gcc tca aag tta ttt gat gca aaa acc Gly Ser Lys Ser Phe Ala Thr Ala Ser Lys Leu Phe Asp Ala Lys Thr 20 25 30 96 cgc agc gta ctg atg ctc tac gcc tgg tgc cgc cat tgt gac gat Arg Ser Val Leu Met Leu Tyr Ala Trp Cys Arg His Cys Asp Asp 35 40 45 144 gtt att gac gat cag acg ctg ggc ttt cag gcc cgg cag cct gcc tta Val Ile Asp Asp Gln Thr Leu Gly Phe Gln Ala Arg Gln Pro Ala Leu 50 55 60 192 caa acg ccc gaa caa cgt ctg atg caa ctt gag atg aaa acg cgc cag Gln Thr Pro Glu Gln Arg Leu Met Gln Leu Glu Met Lys Thr Arg Gln 65 70 75 80 240 gcc tat gca gga tcg cag atg cac gaa ccg gcg ttt gcg gct ttt cag Ala Tyr Ala Gly Ser Gln Met His Glu Pro Ala Phe Ala Ala Phe Gln 85 90 95 288 gaa gtg gct atg gct cat gat atc gcc ccg gct tac gcg ttt gat cat Glu Val Ala Met Ala His Asp Ile Ala Pro Ala Tyr Ala Phe Asp His 100 105 110 336 ctg gaa ggc ttc gcc atg gat gta cgc gaa gcg caa tac agc caa ctg Leu Glu Gly Phe Ala Met Asp Val Arg Glu Ala Gln Tyr Ser Gln Leu 115 120 125 384 t gat acg ctg cgc tat tgc tat cac gtt gca ggc gtt gtc ggc ttg ksp Asp Thr Leu Arg Tyr Cys Tyr His Val Ala Gly Val Val Gly Leu 130 140 432 atg atg gcg caa atc atg ggc gtg cgg gat aac gcc acg ctg gac cgc . Met Met Ala Gln Ile Met Gly Val Arg Asp Asn Ala Thr Leu Asp Arg 145 150 155 480 gcc tgt gac ctt ggg ctg gca ttt cag ttg acc aat att gct cgc gat Ala Cys Asp Leu Gly Leu Ala Phe Gln Leu Thr Asn Ile Ala Arg Asp 165 170 175 528

att gtg gac gat gcg cat gcg ggc cgc tgt tat ctg ccg gca agc tgg Ile Val Asp Asp Ala His Ala Gly Arg Cys Tyr Leu Pro Ala Ser Trp 180 185 190 576 ctg gag cat gaa ggt ctg aac aaa gag aat tat gcg gca cct gaa aac Leu Glu His Glu Gly Leu Asn Lys Glu Asn Tyr Ala Ala Pro Glu Asn 195 200 205 624 cgt cag gcg ctg agc cgt atc gcc cgt cgt ttg gtg cag gaa gca gaa 672 Seite 105

# 

<210> 70

<211> 309

<212> PRT

<213> Erwinia uredovora

<400> 70

Met Asn Asn Pro Ser Leu Leu Asn His Ala Val Glu Thr Met Ala Val 1 10 15

Gly Ser Lys Ser Phe Ala Thr Ala Ser Lys Leu Phe Asp Ala Lys Thr 20 25 30

rg Arg Ser Val Leu Met Leu Tyr Ala Trp Cys Arg His Cys Asp Asp 35 40 45

Val Ile Asp Asp Gln Thr Leu Gly Phe Gln Ala Arg Gln Pro Ala Leu 50 60

Gln Thr Pro Glu Gln Arg Leu Met Gln Leu Glu Met Lys Thr Arg Gln 65 70 75 80

Ala Tyr Ala Gly Ser Gln Met His Glu Pro Ala Phe Ala Ala Phe Gln
85 90 95

Glu Val Ala Met Ala His Asp Ile Ala Pro Ala Tyr Ala Phe Asp His 100 105 110 Leu Glu Gly Phe Ala Met Asp Val Arg Glu Ala Gln Tyr Ser Gln Leu

Asp Asp Thr Leu Arg Tyr Cys Tyr His Val Ala Gly Val Val Gly Leu

Met Met Ala Gln Ile Met Gly Val Arg Asp Asn Ala Thr Leu Asp Arg
160

Ala Cys Asp Leu Gly Leu Ala Phe Gln Leu Thr Asn Ile Ala Arg Asp
180 Ala His Ala Gly Arg Cys Tyr Leu Pro Ala Ser Trp

Leu Glu His Glu Gly Leu Asn Lys Glu Asn Tyr Ala Ala Pro Glu Asn 195 200 205

g Gln Ala Leu Ser Arg Ile Ala Arg Arg Leu Vàl Gln Glu Ala Glu 210 220

Pro Tyr Tyr Leu Ser Ala Thr Ala Gly Leu Ala Gly Leu Pro Leu Arg 225 230 235 240

Ser Ala Trp Ala Ile Ala Thr Ala Lys Gln Val Tyr Arg Lys Ile Gly 245 250 . 255

Val Lys Val Glu Gln Ala Gly Gln Gln Ala Trp Asp Gln Arg Gln Ser 260 265 270

Thr Thr Thr Pro Glu Lys Leu Thr Leu Leu Leu Ala Ala Ser Gly Gln 275 280 285

Ala Leu Thr Ser Arg Met Arg Ala His Pro Pro Arg Pro Ala His Leu 290 295 300

rp Gln Arg Pro Leu 305

<210> 71

<211> 1479

<212> DNA

<213> Erwinia uredovora

<220>

<221> CDS

<222> (1)..(1479)

		_	_														
	ato	> 7 aaa Lys	cca	act Thr	acg Thr 5	gta Val	att Ile	ggt Gly	gca Ala	ggc Gly 10	ttc Phe	ggt Gly	ggc Gly	ctg Leu	gca Ala 15	ctg Leu	48
	gca Ala	att Ile	cgt Arg	cta Leu 20	caa Gln	gct Ala	gcg Ala	ggg Gly	atc Ile 25	ccc Pro	gtc Val	tta Leu	ctg Leu	ctt Leu 30	gaa Glu	caa Gln	96
	cgt Arg	gat Asp	aaa Lys 35	ccc Pro	ggc Gly	ggt Gly	cgg Arg	gct Ala 40	tat Tyr	gtc Val	tac Tyr	gag Glu	gat Asp 45	cag Gln	ggg Gly	ttt Phe	144
	acc Thr	ttt Phe 50	gat Asp	gca Ala	ggc Gly	ccg Pro	acg Thr 55	gtt Val	atc Ile	acc Thr	gat Asp	ccc Pro 60	agt Ser	gcc Ala	att Ile	gaa Glu	192
	gaa Glu	ctg Leu	ttt Phe	gca Ala	ctg Leu	gca Ala 70	gga Gly	aaa Lys	cag Gln	tta Leu	aaa Lys 75	gag Glu	tat Tyr	gtc Val	gaa Glu	ctg Leu 80	240
	.g Leu	ccg Pro	gtt Val	acg Thr	ccg Pro 85	ttt Phe	tac Tyr	cgc Arg	ctg Leu	tgt Cys 90	tgg Trp	gag Glu	tca Ser	ggg Gly	aag Lys 95	gtc Val	288
	ttt Phe	aat Asn	tac Tyr	gat Asp 100	aac Asn	gat Asp	caa Gln	acc Thr	cgg Arg 105	ctc Leu	gaa Glu	gcg Ala	cag Gln	att Ile 110	Gln	cag Gln	336
	ttt Phe	aat Asn	ccc Pro 115	Arg	gat Asp	gtc Val	gaa Glu	ggt Gly 120	tat Tyr	cgt Arg	cag Gln	ttt Phe	ctg Leu 125	Asp	tat Tyr	tca Ser	384
	cgc Arg	gcg Ala 130	٧a٦	ttt Phe	aaa Lys	gaa Glu	ggc Gly 135	tat Tyr	ctà Leu	aag Lys	ctc Leu	ggt Gly 140	Thr	gtc Val	cct Pro	ttt Phe	432
	tta Leu 145	Ser	ttc Phe	aga Arg	gac Asp	atg Met 150	Leu	cgc Arg	gcc Ala	gca Ala	cct Pro 155	Gln	ctg Leu	gcg Ala	aaa Lys	ctg Leu 160	480
6	cag	gca Ala	tgg Trp	aga Arg	agc Ser 165	' Val	tac Tyr	agt Ser	aag Lys	gtt Val 170	Ala	agt Ser	tac Tyr	ato Ile	gaa Glu 175	gat Asp	528
	gaa Glu	cat His	ctg Leu	cgc Arg 180	G]n	gcg Ala	ttt Phe	: Ser	ttc Phe 185	His	tcg Ser	ctg Leu	ttg Lei	gtg Val 190	Gly	ggc Gly	576
•				Āla					Tyr			ille		Ālā		gag Glu	624
	cgt Arg	gag Glu 210	ı Trp	ggo Gly	gto Val	tgg Trp	ttt Phe 215	Pro	cgt Arg	ggc	ggc Gly	aco Thr 220	· Gly	gca / Ala	a tta a Lei	a gtt u Val	672
	cag Glr 225	ı Gly	atg Met	ata Ile	aag Lys	cto Lei 230	ı Phe	cag Glr	gat Asp	ct <u>c</u> Leu	ggt Gly 235	/ Gly	gaa Glu	a gte u Va	c gtg l Va	tta Leu 240	720
	aac Asr	gco Ala	aga Arg	a gto g Val	ago Ser 245	· His	ato Mei	g gaa E Glu	a acg I Thr	• Thr 250	· Gly	/ Ast	aaq 1 Ly:	g at	t gaa e Gli 25	a gcc u Ala 5	768

	gtg Val	cat His	tta Leu	gag Glu 260	gac Asp	ggt Gly	cgc Arg	agg Arg	ttc Phe 265	ctg Leu	acg Thr	caa Gln	gcc Ala	gtc Val 270	gcg Ala	tca Ser		816
	aat Asn	gca Ala	gat Asp 275	gtg Val	gtt Val	cat His	acc Thr	tat Tyr 280	cgc Arg	gac Asp	ctg Leu	tta Leu	agc Ser 285	cag Gln	cac His	cct Pro		864
	gcc Ala	gcg Ala 290	gtt Val	aag Lys	cag Gln	tcc Ser	aac Asn 295	aaa Lys	ctg Leu	cag Gln	act Thr	aag Lys 300	cgc Arg	atg Met	agt Ser	aac Asn		912
	tct Ser 305	ctg Leu	ttt Phe	gtg Val	ctc Leu	tat Tyr 310	ttt Phe	ggt Gly	ttg Leu	aat Asn	cac His 315	cat His	cat His	gat Asp	Gln	ctc Leu 320		960
	gcg Ala	cat His	cac His	acg Thr	gtt Val 325	tgt Cys	ttc Phe	ggc Gly	ccg Pro	cgt Arg 330	tac Tyr	cgc Arg	gag Glu	ctg Leu	att Ile 335	gac Asp		1008
4	gaa Glu	att Ile	ttt Phe	aat Asn 340	cat His	gat Asp	ggc Gly	ctc Leu	gca Ala 345	gag Glu	gac Asp	ttc Phe	tca Ser	ctt Leu 350	tat Tyr	ctg Leu		1056
	C His	gcg Ala	ccc Pro 355	tgt Cys	gtc Val	acg Thr	gat Asp	tcg Ser 360	tca Ser	ctg Leu	gcg Ala	cct Pro	gaa Glu 365	ggt Gly	tgc Cys	ggc Gly		1104
	agt Ser	tac Tyr 370	tat Tyr	gtg Val	ttg Leu	gcg Ala	ccg Pro 375	gtg Val	ccg Pro	cat His	tta Leu	ggc Gly 380	Thr	gcg Ala	aac Asn	ctc Leu		1152
	gac Asp 385	tgg Trp	acg Thr	gtt Val	gag Glu	ggg Gly 390	cca Pro	aaa Lys	cta Leu	cgc Arg	gac Asp 395	Arg	att Ile	ttt Phe	gcg Ala	tac Tyr 400		1200
	ctt Leu	gag Glu	cag Gln	cat His	tac Tyr 405	Met	cct Pro	ggc Gly	tta Leu	cgg Arg 410	Ser	cag Gln	ctg Leu	gtc Val	acg Thr 415	cac His		1248
	cgg Arg	atg Met	ttt Phe	acg Thr 420	Pro	ttt Phe	gat Asp	ttt Phe	cgc Arg 425	Asp	cag Gln	ctt Leu	aat Asn	gcc Ala 430	Tyr	cat His		1296
				Phe					٧al					Āla		ttt Phe		1344
	cgg Arg	ccg Pro 450	His	aac Asr	cgc Arg	gat Asp	aaa Lys 455	Thr	att	act Thr	aat Asn	cto Leu 460	ı Tyr	ctg Leu	g gto I Val	ggc		1392
	gca Ala 465	Gly	acg Thr	cat His	ccc Pro	ggc Gly 470	' Ala	ggc Gly	att Ile	cct Pro	ggo Gly 475	/ Val	ato Ile	ggo Gly	tcg / Ser	gca Ala 480	: .	1440
	aaa Lys	gcg Ala	aca Thr	gca Ala	a ggt a Gly 485	/ Lei	ı atçı Met	ctg Lei	gag Glu	gat Asp 490	) Lei	g ata u Ile	tga e	a				1479

<210> 72

<211> 492

<212> PRT

<400> 72

Met Lys Pro Thr Thr Val Ile Gly Ala Gly Phe Gly Gly Leu Ala Leu 1 10 15

Ala Ile Arg Leu Gln Ala Ala Gly Ile Pro Val Leu Leu Glu Gln 20 25 30

Arg Asp Lys Pro Gly Gly Arg Ala Tyr Val Tyr Glu Asp Gln Gly Phe 35 40 45

Thr Phe Asp Ala Gly Pro Thr Val Ile Thr Asp Pro Ser Ala Ile Glu 50 60

Glu Leu Phe Ala Leu Ala Gly Lys Gln Leu Lys Glu Tyr Val Glu Leu 70 75 80

Leu Pro Val Thr Pro Phe Tyr Arg Leu Cys Trp Glu Ser Gly Lys Val 85 90 95

Phe Asn Tyr Asp Asn Asp Gln Thr Arg Leu Glu Ala Gln Ile Gln Gln 100 105 110

Phe Asn Pro Arg Asp Val Glu Gly Tyr Arg Gln Phe Leu Asp Tyr Ser 115 120 125

Arg Ala Val Phe Lys Glu Gly Tyr Leu Lys Leu Gly Thr Val Pro Phe 130 140

Leu Ser Phe Arg Asp Met Leu Arg Ala Ala Pro Gln Leu Ala Lys Leu 145 150 155 160

In Ala Trp Arg Ser Val Tyr Ser Lys Val Ala Ser Tyr Ile Glu Asp 165 170 175

Glu His Leu Arg Gln Ala Phe Ser Phe His Ser Leu Leu Val Gly Gly 180 185 190

Asn Pro Phe Ala Thr Ser Ser Ile Tyr Thr Leu Ile His Ala Leu Glu 195 200 · 205

Arg Glu Trp Gly Val Trp Phe Pro Arg Gly Gly Thr Gly Ala Leu Val 210 215 220

Gln Gly Met Ile Lys Leu Phe Gln Asp Leu Gly Glu Val Val Leu 225 230 235 240

Asn Ala Arg Val Ser His Met Glu Thr Thr Gly Asn Lys Ile Glu Ala 245 250 255 Seite 110

Val His Leu Glu Asp Gly Arg Arg Phe Leu Thr Gln Ala Val Ala Ser 260 265 270

Asn Ala Asp Val Val His Thr Tyr Arg Asp Leu Leu Ser Gln His Pro 275 280 285

Ala Ala Val Lys Gln Ser Asn Lys Leu Gln Thr Lys Arg Met Ser Asn 290 295 300

Ser Leu Phe Val Leu Tyr Phe Gly Leu Asn His His His Asp Gln Leu 305 310 315 320

Ala His His Thr Val Cys Phe Gly Pro Arg Tyr Arg Glu Leu Ile Asp 325 330 335

Glu Ile Phe Asn His Asp Gly Leu Ala Glu Asp Phe Ser Leu Tyr Leu 340 345 350

His Ala Pro Cys Val Thr Asp Ser Ser Leu Ala Pro Glu Gly Cys Gly 355 360 365

Ser Tyr Tyr Val Leu Ala Pro Val Pro His Leu Gly Thr Ala Asn Leu 370 375 380

Asp Trp Thr Val Glu Gly Pro Lys Leu Arg Asp Arg Ile Phe Ala Tyr 385 390 395 400

Leu Glu Gln His Tyr Met Pro Gly Leu Arg Ser Gln Leu Val Thr His 405 410 415

Arg Met Phe Thr Pro Phe Asp Phe Arg Asp Gln Leu Asn Ala Tyr His 420 425 430

ly Ser Ala Phe Ser Val Glu Pro Val Leu Thr Gln Ser Ala Trp Phe 435 . 440 445

Arg Pro His Asn Arg Asp Lys Thr Ile Thr Asn Leu Tyr Leu Val Gly
450 460

Ala Gly Thr His Pro Gly Ala Gly Ile Pro Gly Val Ile Gly Ser Ala 465 470 475 480

Lys Ala Thr Ala Gly Leu Met Leu Glu Asp Leu Ile 485 490

<210> 73

<211> 1725

<212> DNA

•																	
	<220>																
	<221>	C	S					•									
	<222>	(1	L)	(172	5)												
	<223>	•															
	<400> atg @ Met # 1	ict .	tct	tcc Ser	act Thr 5	tgt Cys	tta Leu	att Ile	HIS	tct Ser 10	tcc Ser	tct Ser	ttt Phe	ggg Gly	gtt Val 15	gga Gly	48
	gga a Gly i	aag Lys	aaa Lys	gtg Val 20	aag Lys	atg Met	aac Asn	acg Thr	atg Met 25	att Ile	cga Arg	tcg Ser	aag Lys	ttg Leu 30	ttt Phe	tca Ser	96
	t e	cgg Arg	tcg Ser 35	gct Ala	ttg Leu	gac Asp	act Thr	aag Lys 40	gtg Val	tct Ser	gat Asp	atg Met	agc Ser 45	gtc Val	aat Asn	gct Ala	144
	Pro	aaa Lys 50	gga Gly	ttg Leu	ttt Phe	cca Pro	cca Pro 55	gag Glu	cct Pro	gag Glu	cac His	tac Tyr 60	agg Arg	gġg Gly	cca Pro	aag Lys	192
	ctt	222	gtg Val	gct Ala	atc Ile	att Ile 70	gga Gly	gct Ala	ggg Gly	ctc Leu	gct Ala 75	ggc Gly	atg Met	tca Ser	act Thr	gca Ala 80	240
	gtg Val	gag Glu	ctt Leu	ttg Leu	gat Asp 85	caa Gln	ggg Gly	cat His	gag Glu	gtt Val 90	gac Asp	ata Ile	tat Tyr	gaa Glu	tcc Ser 95	aga Arg	288
	caa Gln	ttt Phe	att Ile	ggt Gly 100	ggt Gly	aaa Lys	gtc Val	ggt Gly	tct Ser 105	ttt Phe	gta Val	gat Asp	aag Lys	cgt Arg 110	GIY	aac Asn	336
	cat His	att Ile	gaa Glu 115	Met	gga Gly	ctc Leu	cat His	gtg Val 120	Phe	ttt Phe	ggt Gly	tgc Cys	tat Tyr 125	ASN	aat Asr	ctt Leu	384
	ne ne	aga Arg 130	Leu	atg Met	aaa Lys	aag Lys	gta Val 135	Gly	gca Ala	gat Asp	gaa Glu	aat Asn 140	Leu	ctg Leu	gto Va	aag Lys	432
	gat Asp 145	His	act Thr	cat His	acc Thr	ttt Phe 150	· va i	aac Asn	cga	ggt	gga Gly 155	GIL	att Ile	ggt Gly	gaa / Gl	ctt Leu 160	480
	gat Asp	ttc Phe	cga	cti Leu	ccg Pro 165	) Met	ggt Gly	gca Ala	cca Pro	tta Leu 170	i His	t ggt s Gly	ati / Ile	cg1 Arg	g Ala 17	a ttt a Phe 5	528
	cta Leu	aca Thr	act Thr	aat Asr 180	្ធ Gli	a cto n Lei	aag Lys	g cct s Pro	tat Tyr 185	. Ast	aaa Lys	a gca s Ala	a agg	g aat g Asi 190	n Al	t gtg a Val	576
	gct Ala	ctt Lei	gco Ala 19:	a Lei	t age	c cca r Pro	a gti o Va	t gta l Va 200	i Arg	t gci g Ala	t cti	t at u Il	t ga e As 20	b bl	a aa o As	t ggt n Gly	624
	gca	ato	g ca	g ga	t at	a agg	g aa	c tta	a ga		t at		c tt	t tc	t ga	t tgg	672

Seite 112

Ala	Met 210	Gln	Asp	Ile	Arg	Asn 215	Leu	Asp	Asn	Ile	Ser 220	Phe	Ser	Asp	Trp	
ttc Phe 225	tta Leu	tcc Ser	aaa Lys	ggc Gly	ggt Gly 230	acc Thr	cgc Arg	atg Met	agc Ser	atc Ile 235	caa Gln	agg Arg	atg Met	tgg Trp	gat Asp 240	720
cca Pro	gtt Val	gct Ala	tat Tyr	gcc Ala 245	ctc Leu	gga Gly	ttt Phe	att Ile	gac Asp 250	tgt Cys	gat Asp	aat Asn	atc Ile	agt Ser 255	gcc Ala	768
cgt Arg	tgt Cys	atg Met	ctt Leu 260	act Thr	ata Ile	ttt Phe	tct Ser	cta Leu 265	ttt Phe	gct Ala	act Thr	aag Lys	aca Thr 270	gaa Glu	gct Ala	816
tct Ser	ctg Leu	ttg Leu 275	cgt Arg	atg Met	ttg Leu	aag Lys	ggt Gly 280	tcg Ser	cct Pro	gat Asp	gtt Val	tac Tyr 285	tta Leu	agc Ser	ggt Gly	864
cct Pro	ata Ile 290	aga Arg	aag Lys	tat Tyr	att Ile	aca Thr 295	gat Asp	aaa Lys	ggt Gly	gga Gly	agg Arg 300	ttt Phe	cac His	cta Leu	agg Arg	912
g	ggg Gly	tgt Cys	aga Arg	gag Glu	ata Ile 310	ctt Leu	tat Tyr	gat Asp	gaa Glu	cta Leu 315	tca Ser	aat Asn	ggc Gly	gac Asp	aca Thr 320	960
tat Tyr	atc Ile	aca Thr	ggc Gly	att Ile 325	gca Ala	atg Met	tcg Ser	aag Lys	gct Ala 330	acc Thr	aat Asn	aaa Lys	aaa Lys	ctt Leu 335	gtġ Val	1008
aaa Lys	gct Ala	gac Asp	gtg Val 340	tat Tyr	gtt Val	gca Ala	gca Ala	tgt Cys 345	gat Asp	gtt Val	cct Pro	gga Gly	ata Ile 350	aaa Lys	agg Arg	1056
ttg Leu	atc Ile	cca Pro 355	tcg Ser	gag Glu	tgg Trp	aga Arg	gaa Glu 360	tgg Trp	gat Asp	cta Leu	ttt Phe	gac Asp 365	aat Asn	atc Ile	tat Tyr	1104
aaa Lys	cta Leu 370	gtt Val	gga Gly	gtt Val	cca Pro	gtt Val 375	gtc Val	act Thr	gtt Val	cag Gln	ctt Leu 380	agg Arg	tac Tyr	aat Asn	ggt Gly	1152
tgg Trp 385	gtg Val	aca Thr	gag Glu	atg Met	caa Gln 390	Asp	ctg Leu	gaa Glu	aaa Lys	tca Ser 395	agg Arg	cag Gln	ttg Leu	aga Arg	gct Ala 400	1200
a	gta Val	gga Gly	ttg Leu	gat Asp 405	aat Asn	ctt Leu	ctt Leu	tat Tyr	act Thr 410	cca Pro	gat Asp	gca Ala	gac Asp	ttt Phe 415	Ser	1248
tgt Cys	ttt Phe	tct Ser	gat Asp 420	ctt Leu	gca Ala	ctc Leu	tcg Ser	tcg Ser 425	cct Pro	gaa Glu	gat Asp	tat Tyr	tat Tyr 430	Ile	gaa Glu	1296
gga Gly	caa Gln	ggg Gly 435	Ser	cta Leu	ata Ile	cag Gln	gct Ala 440	٧a٦	ctc Leu	acg Thr	cca Pro	ggg Gly 445	gat Asp	cca Pro	tac Tyr	1344
atg Met	CCC Pro 450	Leu	cct Pro	aat Asn	gat Asp	gca Ala 455	Ile	ata Ile	gaa Glu	aga Arg	gtt Val 460	Arg	aaa Lys	cag Gln	gtt Val	1392
ttg Leu 465	ASP	tta Leu	ttc Phe	cca Pro	tcc Ser 470	Ser	caa Gln	ggc Gly	ctg Leu	gaa Glu 475	gtt Val	cta Leu	tgg Trp	tct Ser	tcg Ser 480	1440
gtg	gtt	aaa	atc	gga	caa	tcc	cta	tat		gag te 1		cct	gga	aag	gac	1488

ValValLysIleGly<br/>485GlnSerLeuTyrArg<br/>ArgGluGly<br/>GluProGly<br/>495Lys<br/>AspAspCcattc<br/>Pheaga<br/>Arg<br/>Soocca<br/>Glnaaa<br/>Asp<br/>Glnaca<br/>Lyscca<br/>Thr<br/>Soogta<br/>Asp<br/>Thr<br/>Soocca<br/>Asp<br/>Soogta<br/>Asp<br/>Thr<br/>Sooaaa<br/>Asp<br/>Tyr<br/>Sooaaa<br/>Asp<br/>Tyr<br/>Thr<br/>Sooaca<br/>Asp<br/>Tyr<br/>Tyr<br/>Thr<br/>Sooata<br/>Asp<br/>Tyr<br/>Tyr<br/>Thr<br/>Sooata<br/>Asp<br/>Tyr<br/>Tyr<br/>The<br/>Asp<br/>Sooata<br/>Asp<br/>Tyr<br/>Thr<br/>Asp<br/>Sooata<br/>Tyr<br/>Thr<br/>Thr<br/>Asp<br/>Sooata<br/>Tyr<br/>Thr<br/>Thr<br/>Asp<br/>Sooata<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooata<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Sooaga<br/>Tyr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<br/>Thr<b

10> 74

<211> 574

<212> PRT

<213> Narcissus pseudonarcissus

<400> 74

Met Ala Ser Ser Thr Cys Leu Ile His Ser Ser Ser Phe Gly Val Gly 10 10

Gly Lys Lys Val Lys Met Asn Thr Met Ile Arg Ser Lys Leu Phe Ser 20 25 30

Ile Arg Ser Ala Leu Asp Thr Lys Val Ser Asp Met Ser Val Asn Ala 35 40 45

ro Lys Gly Leu Phe Pro Pro Glu Pro Glu His Tyr Arg Gly Pro Lys
50 55 60

Leu Lys Val Ala Ile Ile Gly Ala Gly Leu Ala Gly Met Ser Thr Ala 65 70 75 80

Val Glu Leu Leu Asp Gln Gly His Glu Val Asp Ile Tyr Glu Ser Arg 85 90 95

Gln Phe Ile Gly Gly Lys Val Gly Ser Phe Val Asp Lys Arg Gly Asn 100 105 110

His Ile Glu Met Gly Leu His Val Phe Phe Gly Cys Tyr Asn Asn Leu 115 120 125 O4sequ.txt
Phe Arg Leu Met Lys Lys Val Gly Ala Asp Glu Asn Leu Leu Val Lys
130 140 Asp His Thr His Thr Phe Val Asn Arg Gly Glu Ile Gly Glu Leu 145 : 150 155 160 Asp Phe Arg Leu Pro Met Gly Ala Pro Leu His Gly Ile Arg Ala Phe 165 170 175 Leu Thr Thr Asn Gln Leu Lys Pro Tyr Asp Lys Ala Arg Asn Ala Val 180 185 190 Ala Leu Ala Leu Ser Pro Val Val Arg Ala Leu Ile Asp Pro Asn Gly 195 200 205 Ala Met Gln Asp Ile Arg Asn Leu Asp Asn Ile Ser Phe Ser Asp Trp 210 215 220 Leu Ser Lys Gly Gly Thr Arg Met Ser Ile Gln Arg Met Trp Asp 230 240 Pro Val Ala Tyr Ala Leu Gly Phe Ile Asp Cys Asp Asn Ile Ser Ala 245 250 255 Arg Cys Met Leu Thr Ile Phe Ser Leu Phe Ala Thr Lys Thr Glu Ala 260 265 270 Ser Leu Leu Arg Met Leu Lys Gly Ser Pro Asp Val Tyr Leu Ser Gly 275 280 285 Pro Ile Arg Lys Tyr Ile Thr Asp Lys Gly Gly Arg Phe His Leu Arg 290 295 300 Trp Gly Cys Arg Glu Ile Leu Tyr Asp Glu Leu Ser Asn Gly Asp Thr 305 310 315 320 yr Ile Thr Gly Ile Ala Met Ser Lys Ala Thr Asn Lys Lys Leu Val 325 330 335 Lys Ala Asp Val Tyr Val Ala Ala Cys Asp Val Pro Gly Ile Lys Arg 340 345 350 Leu Ile Pro Ser Glu Trp Arg Glu Trp Asp Leu Phe Asp Asn Ile Tyr 355 360 365 Lys Leu Val Gly Val Pro Val Val Thr Val Gln Leu Arg Tyr Asn Gly 370 375 380 Trp Val Thr Glu Met Gln Asp Leu Glu Lys Ser Arg Gln Leu Arg Ala 385 390 395 400 04sequ.txt
Ala Val Gly Leu Asp Asn Leu Leu Tyr Thr Pro Asp Ala Asp Phe Ser
405
410
415

Cys Phe Ser Asp Leu Ala Leu Ser Ser Pro Glu Asp Tyr Tyr Ile Glu 420 425 430

Gly Gln Gly Ser Leu Ile Gln Ala Val Leu Thr Pro Gly Asp Pro Tyr 445 445

Met Pro Leu Pro Asn Asp Ala Ile Ile Glu Arg Val Arg Lys Gln Val 450 460

Leu Asp Leu Phe Pro Ser Ser Gln Gly Leu Glu Val Leu Trp Ser Ser 480

Val Val Lys Ile Gly Gln Ser Leu Tyr Arg Glu Gly Pro Gly Lys Asp 485 490 495

Phe Arg Pro Asp Gln Lys Thr Pro Val Lys Asn Phe Phe Leu Ala 500 500

Gly Ser Tyr Thr Lys Gln Asp Tyr Ile Asp Ser Met Glu Gly Ala Thr 515 520 525

Leu Ser Gly Arg Gln Ala Ala Ala Tyr Ile Cys Ser Ala Gly Glu Asp 530 535

Leu Ala Ala Leu Arg Lys Lys Ile Ala Ala Asp His Pro Glu Gln Leu 545 550 560

Ile Asn Lys Asp Ser Asn Val Ser Asp Glu Leu Ser Leu Val 565 570

<210> 75

211> 1848

212> DNA

<213> Lycopersicon esculentum

<220>

<221> CDS

<222> (1)..(1848)

<223>

<400> 75
atg tgt acc ttg agt ttt atg tat cct aat tca ctt ctt gat ggt acc
Met Cys Thr Leu Ser Phe Met Tyr Pro Asn Ser Leu Leu Asp Gly Thr
10
15
Seite 116

# O4sequ.txt

aga agt tot tgt ttt gac cot ttg ata att gga aat tgt act gat cag	4
Arg Ser Ser Cys Phe Asp Pro Leu Ile Ile Gly Asn Cys Thr Asp Gln 35 40 45	
cag cag ctt tgt ggc ttg agt tgg ggg gtg gac aag gct aag gga aga 19 Gln Gln Leu Cys Gly Leu Ser Trp Gly Val Asp Lys Ala Lys Gly Arg 50 55 60	12
aga ggg ggt act gtt tcc aat ttg aaa gca gtt gta gat gta gac aaa 24 Arg Gly Gly Thr Val Ser Asn Leu Lys Ala Val Val Asp Val Asp Lys 65 70 75 80	10
aga gtg gag agc tat ggc agt agt gat gta gaa gga aat gag agt ggc 28 Arg Val Glu Ser Tyr Gly Ser Ser Asp Val Glu Gly Asn Glu Ser Gly 85 90 95	88
agc tat gat gcc att gtt ata ggt tca gga ata ggt gga ttg gtg gca Ser Tyr Asp Ala Ile Val Ile Gly Ser Gly Ile Gly Gly Leu Val Ala 100 105 110	36
g acg cag ctg gcg gtt aag gga gct aag gtt tta gtt ctg gag aag Ala Thr Gln Leu Ala Val Lys Gly Ala Lys Val Leu Val Leu Glu Lys 115 120 125	84
tat gtt att cct ggt gga agc tct ggc ttt tac gag agg gat ggt tat 4 Tyr Val Ile Pro Gly Gly Ser Ser Gly Phe Tyr Glu Arg Asp Gly Tyr 130 135 140	32
aag ttt gat gtt ggt tca tca gtg atg ttt gga ttc agt gat aag gga Lys Phe Asp Val Gly Ser Ser Val Met Phe Gly Phe Ser Asp Lys Gly 145 150 160	180
aac ctc aat tta att act caa gca ttg gca gca gta gga cgt aaa tta Asn Leu Asn Leu Ile Thr Gln Ala Leu Ala Ala Val Gly Arg Lys Leu 165 170 175	528
gaa gtt ata cct gac cca aca act gta cat ttc cac ctg cca aat gac Glu Val Ile Pro Asp Pro Thr Thr Val His Phe His Leu Pro Asn Asp 180 185 190	576
ctt tct gtt cgt ata cac cga gag tat gat gac ttc att gaa gag ctt eu Ser Val Arg Ile His Arg Glu Tyr Asp Asp Phe Ile Glu Glu Leu 195 200 205	624
gtg agt aaa ttt cca cat gaa aag gaa ggg att atc aaa ttt tac agt Val Ser Lys Phe Pro His Glu Lys Glu Gly Ile Ile Lys Phe Tyr Ser 210 215 220	672
gaa tgc tgg aag atc ttt aat tct ctg aat tca ttg gaa ctg aag tct Glu Cys Trp Lys Ile Phe Asn Ser Leu Asn Ser Leu Glu Leu Lys Ser 225 230 235 240	720
ttg gag gaa ccc atc tac ctt ttt ggc cag ttc ttt aag aag ccc ctt Leu Glu Glu Pro Ile Tyr Leu Phe Gly Gln Phe Phe Lys Lys Pro Leu 245 250 255	768
gaa tgc ttg act ctt gcc tac tat ttg ccc cag aat gct ggt agc atc Glu Cys Leu Thr Leu Ala Tyr Tyr Leu Pro Gln Asn Ala Gly Ser Ile 260 265 270	816
gct cgg aag tat ata aga gat cct ggg ttg ctg tct ttt ata gat gca Ala Arg Lys Tyr Ile Arg Asp Pro Gly Leu Leu Ser Phe Ile Asp Ala 275 280 285 Seite 117	864

	gag tg Glu Cy 29	s Phe	atc Ile	gtg Val	ser	aca Thr 295	gtt Val	aat Asn	gca Ala	tta Leu	caa Gln 300		cca Pro	atg Met	ato	2	912
	aat gc Asn Al 305		atg Met	gtt Val	cta Leu 310	tgt Cys	gac Asp	aga Arg	cat His	ttt Phe 315	ggc Gly	gga Gly	atc Ile	aac Asn	tad Tyl 320	= r 0	960
	ccc gt Pro Va	t ggt il Gly	gga Gly	gtt Val 325	ggc Gly	gag Glu	atc Ile	gcc Ala	aaa Lys 330		tta Leu	gca Ala	aaa Lys	ggc Gly 335		g u	1008
	gat ga Asp As	at cad sp His	gga Gly 340	agt Ser	cag Gln	ata Ile	ctt Leu	tat Tyr 345	Ai y	gca Ala	aat Asr	gtt Val	aca Thr 350		at Il	c e	1056
	att t	tg gae eu Asj 35	aat Asn		aaa Lys	gct Ala	gtg Val 360	gga Gly	gto Val	aag Lys	cti Le	t tct u ser 365	gac Asp	ggg	g ag / Ar	ig 'g	1104
	aag t	tt ta he Ty		aaa Lys	acc Thr	ata Ile 375	gta Val	+ + c (	g aat r Asi	gc1 n Ala	t ac a Th 38		tgg Trp	gat Asp	t ac	it ir	1152
	t g	70 ga aa ily Ly	g ct1 s Lei	t tta u Leu	aaa Lys 390	gct	. 42	aa a As	t ct n Le	g cca u Pre 39	о су	a ga s Gl	a gaa u Glu	a gaa u Gl	a aa u A: 40	at . sn 00	1200
	385 ttc ( Phe (	ag aa In Ly	a gc s Al	t ta a Ty: 40	t gta		a gc	a cc a Pr	t tc o Se 41		t ct e Le	t tc eu Se	t att	t ca e Hi 41	_	tg et	1248
	gga (	gtt aa Val Ly	aa gc /s Al 42	a ga a As		a ct	c cc u Pr	a cc o Pr 42	O AS	c ac p Th	a ga ır As	at tg sp Cy	t ca 's Hi 43		t t s P	tt he	1296
	gtc Val	ctc g	ag ga lu As		t tg p Tr	g ac p Th	a aa r As 44	t tt	-a as	ig aa iu Ly	aa c /s P	ca ta ro Ty 44		a ag y Se	gt a er I	ita (le	1344
	ttc Phe	ttg a Leu S	35 gt at er Il	t co le Pr	a ac o Th	a gt ir Va 45	t ct	-t n:	at to sp S	cc to er So		tg g eu A 60	cc co la Pr	a ga o G	aa g lu G	gga Gly	1392
4	cac is	450 cat a His I	tt ci	tt ca eu Hi	ic at	t tt	·+ a	ca a hr T	ca t hr S	د ہے	gc a er I 75	tt g le G	aa ga lu As	at to sp T	- p-	gag Glu 480	1440
		ctc t Leu S	ct c ser P	ro L			at g yr G	aa g lu A	la L	ag a ys L 90	aa g ys G	gag g Glu V	tt g al V		ct la 95	gaa Glu	1488
	agg Arg	att a	[le S	ac 3	aa c	tt g eu G	aa a lu L	.ys <u>ı</u>	ica c Thr 1	tc t eu F	tc o	cca g Pro G	.,, <u> </u>	tt a eu L 10	ag .ys	tca Ser	1536
	tct Ser	att (	ctc t Leu F		ag g ys G	ag g lu V	aic		oct (	cca a Pro l	aag _ys	• • • • •	cac a His A	iga d krg A	cga Arg	tac Tyr	1584
	ctt Leu	gct Ala	515 cgt g Arg A	gat a Asp S	igt g Ser G	ו עוו	cc 1	rat d	gga Gly	cca a	atg Met	cca ( Pro /	cgc g Arg (	gga i	aca Thr	cct Pro	1632
	aag Lys 545	530 gga Gly	ctc ( Leu	ctg g Leu (	aly M	+~ /		ttc Phe	aat Asn	1111	act Thr 555	gct Ala	ata (	gat Asp	ggt Gly	cta Leu 560	1680
	J-T-	•							:	seit	e 11	.8					

tat Tyr	tgt Cys	gtt Val	ggc Gly	gat Asp 565	agt Ser	tgc Cys	ttc Phe	cca Pro	gga Gly 570	caa Gln	ggt Gly	gtt Val	ata Ile	gct Ala 575	gta Val	1728
gcc Ala	ttt Phe	tca Ser	gga Gly 580	gta Val	atg Met	tgc Cys	gct Ala	cat His 585	cgt Arg	gtt Val	gca Ala	gct Ala	gac Asp 590	tta Leu	ggg Gly	1776
ttt Phe	gaa Glu	aaa Lys 595	aaa Lys	tca Ser	gat Asp	gtg Val	ctg Leu 600	gac Asp	agt Ser	gct Ala	ctt Leu	ctt Leu 605	aga Arg	cta Leu	ctt Leu	1824
ggt Gly	tgg Trp 610	tta Leu	agg Arg	aca Thr	cta Leu	gca Ala 615	tga									1848

<210> 76

<211> 615

<212> PRT

13> Lycopersicon esculentum

<400> 76

Met Cys Thr Leu Ser Phe Met Tyr Pro Asn Ser Leu Leu Asp Gly Thr 1 10 15

Cys Lys Thr Val Ala Leu Gly Asp Ser Lys Pro Arg Tyr Asn Lys Gln 20 . 30

Arg Ser Ser Cys Phe Asp Pro Leu Ile Ile Gly Asn Cys Thr Asp Gln 35 40 45

Gln Gln Leu Cys Gly Leu Ser Trp Gly Val Asp Lys Ala Lys Gly Arg 50 60

rg Gly Gly Thr Val Ser Asn Leu Lys Ala Val Val Asp Val Asp Lys 70 75

Arg Val Glu Ser Tyr Glý Ser Ser Asp Val Glu Gly Asn Glu Ser Gly 85 90 95

Ser Tyr Asp Ala Ile Val Ile Gly Ser Gly Ile Gly Gly Leu Val Ala 100 105 110

Ala Thr Gln Leu Ala Val Lys Gly Ala Lys Val Leu Val Leu Glu Lys 115 120 125

Tyr Val Ile Pro Gly Gly Ser Ser Gly Phe Tyr Glu Arg Asp Gly Tyr 130 140

Lys Phe Asp Val Gly Ser Ser Val Met Phe Gly Phe Ser Asp Lys Gly 145 150 160 Seite 119 Asn Leu Asn Leu Ile Thr Gln Ala Leu Ala Ala Val Gly Arg Lys Leu Glu Val Ile Pro Asp Pro Thr Thr Val His Phe His Leu Pro Asn Asp 190 Asn Asp Leu Ser Val Arg Ile His Arg Glu Tyr Asp Asp Phe Ile Glu Glu Leu Val Ser Lys Phe Pro His Glu Lys Glu Gly Ile Ile Lys Phe Tyr Ser Glu Cys Trp Lys Ile Phe Asn Ser Leu Asn Ser Leu Glu Leu Lys Ser 230 Leu Glu Glu Pro Ile Tyr Leu Phe Gly Gln Phe Phe Lys Lys Pro Leu Glu Cys Leu Thr Leu Ala Tyr Tyr Leu Pro Gln Asn Ala Gly Ser Ile Ala Arg Lys Tyr Ile Arg Asp Pro Gly Leu Leu Ser Phe Ile Asp Ala Glu Cys Phe Ile Val Ser Thr Val Asn Ala Leu Gln Thr Pro Met Ile 290 Cys Phe Ile Val Ser Thr Val Asn Ala Leu Gl

Asn Ala Ser Met Val Leu Cys Asp Arg His Phe Gly Gly Ile Asn Tyr 305 310 315 320

Pro Val Gly Gly Val Gly Glu Ile Ala Lys Ser Leu Ala Lys Gly Leu 325 330 335

sp Asp His Gly Ser Gln Ile Leu Tyr Arg Ala Asn Val Thr Ser Ile 340 345 350

Ile Leu Asp Asn Gly Lys Ala Val Gly Val Lys Leu Ser Asp Gly Arg 355 360 365

Lys Phe Tyr Ala Lys Thr Ile Val Ser Asn Ala Thr Arg Trp Asp Thr 370 380

Phe Gly Lys Leu Leu Lys Ala Glu Asn Leu Pro Lys Glu Glu Glu Asn 385 390 395

Phe Gln Lys Ala Tyr Val Lys Ala Pro Ser Phe Leu Ser Ile His Met 405 410 415

Gly Val Lys Ala Asp Val Leu Pro Pro Asp Thr Asp Cys His His Phe 420 430 Seite 120

Val Leu Glu Asp Asp Trp Thr Asn Leu Glu Lys Pro Tyr Gly Ser Ile 435 440 445

Phe Leu Ser Ile Pro Thr Val Leu Asp Ser Ser Leu Ala Pro Glu Gly 450 460

His His Ile Leu His Ile Phe Thr Thr Ser Ser Ile Glu Asp Trp Glu 465 470 480

Gly Leu Ser Pro Lys Asp Tyr Glu Ala Lys Lys Glu Val Val Ala Glu 485 490

Arg Ile Ile Ser Arg Leu Glu Lys Thr Leu Phe Pro Gly Leu Lys Ser 500 505

Ser Ile Leu Phe Lys Glu Val Gly Thr Pro Lys Thr His Arg Arg Tyr 525

Leu Ala Arg Asp Ser Gly Thr Tyr Gly Pro Met Pro Arg Gly Thr Pro 530

Lys Gly Leu Leu Gly Met Pro Phe Asn Thr Thr Ala Ile Asp Gly Leu 545 550 550 560

Tyr Cys Val Gly Asp Ser Cys Phe Pro Gly Gln Gly Val Ile Ala Val 565 570

Ala Phe Ser Gly Val Met Cys Ala His Arg Val Ala Ala Asp Leu Gly 580 580 580

Phe Glu Lys Lys Ser Asp Val Leu Asp Ser Ala Leu Leu Arg Leu Leu 595 600

y Trp Leu Arg Thr Leu Ala 610 615

<210> 77

<211> 1233

<212> DNA

<213> Tagetes erecta

<220>

<221> CDS

<222> (1)..(1233)

<223>

<400 atg Met 1	acc	7 aca Thr	cac His	aaa Lys 5	ctc Leu	ctt Leu	caa Gln	ttc Phe	acc Thr 10	acc Thr	aat Asn	ctc Leu	cca Pro	cca Pro 15	tct Ser	48
tct Ser	tct Ser	tca Ser	atc Ile 20	tct Ser	act Thr	ggc Gly	tgt Cys	tca ser 25	ctc Leu	tcc Ser	ccc Pro	ttc Phe	ttc Phe 30	ctc Leu	aaa Lys	96
tca Ser	tct Ser	tct Ser 35	cat His	tcc Ser	cct Pro	aac Asn	cct Pro 40	cgc Arg	cga Arg	cac His	cgc Arg	cgc Arg 45	tcc Ser	gcc Ala	gta Val	144
tgc Cys	tgc Cys 50	tct Ser	ttc Phe	gcc Ala	tca Ser	ctc Leu 55	gac Asp	tct Ser	gca Ala	aaa Lys	atc Ile 60	aaa Lys	gtc Val	gtt Val	ggc Gly	192
gtc Val 65	ggt Gly	ggt Gly	ggt Gly	ggc Gly	aac Asn 70	aat Asn	gcc Ala	gtt Val	aac Asn	cgc Arg 75	atg Met	att Ile	ggt Gly	agc Ser	ggc Gly 80	240
a	cag Gln	ggt Gly	gtt Val	gat Asp 85	ttt Phe	tac Tyr	gcc Ala	att Ile	aac Asn 90	acg Thr	gac Asp	tca Ser	caa Gln	gcg Ala 95	ctt Leu	288
ctg Leu	caa Gln	tċt Ser	gtt Val 100	gca Ala	cat His	aac Asn	cct Pro	att Ile 105	caa Gln	att Ile	ggg Gly	gag Glu	ctt Leu 110	ttg Leu	act Thr	336
cgt Arg	gga Gly	tta Leu 115	Gly	act Thr	ggt Gly	ggg Gly	aac Asn 120	ccg Pro	ctt Leu	ttg Leu	gga Gly	gaa Glu 125	Gin	gct Ala	gcg Ala	384
gag Glu	gag Glu 130	Ser	aag Lys	gaa Glu	gcg Ala	att Ile 135	ggg Gly	aat Asn	gcg Ala	ctt Leu	aaa Lys 140	Gly	tcg Ser	gat Asp	ctt Leu	432
gtg Val 145	Phe	ata Ile	aca Thr	gca Ala	ggt Gly 150		ggt Gly	ggt Gly	ggg	acg Thr 155		tcg Ser	ggt	gct	gct Ala 160	480
cca Pro	gtt Val	gta Val	gcg Ala	cag Gln 165	Ile	gcg Ala	aaa Lys	gaa Glu	gca Ala 170	Gly	tat Tyr	tta Leu	act Thr	gtt Val 175	ggt Gly	528
al al	gta Val	acg Thr	tac Tyr 180	Pro	tto Phe	agc Ser	· Phe	gaa Glu 185	Gly	cgt Arg	aaa Lys	aga Arg	tca Ser 190	' Va	cag   Gln	576
gcg Ala	j tta Lei	gag Glu 195	ıĀla	att a Ile	gag Glu	j aag i Lys	ctg Leu 200	ı Glr	aag Lys	aac Asr	gtt Val	gad Asp 20:	Thi	tei	t ata u Ile	624
gtg Va	ati   Ile   210	Pro	a aat o Ast	t gad n Asp	cgt Arg	ttig Lei 213	Lei	gat I Asp	att Ile	gc1 Ala	gat a Asp 220	) Glu	a aad 1 Asi	ace Thi	g cct r Pro	672
cti Lei 223	ı Gli	g gar n As <sub>l</sub>	t gct o Ala	t tti a Phe	t cti e Lei 230	រ Leu	gct Ala	gat a Asp	gat Asp	t gta Va 23:	l Lei	cg Ar	c car g Gl	agg nGl	a gtt y Val 240	720
caa Gli	a gga n Gly	a ate	c tc e Se	a gar r Ası 24	p Il	a ati e Ile	t aca e Thi	a ata r Ile	e Pro 250	o Gi	g ctg y Led	g gt u Va	a aa 1 As	t gt n Va 25	g gac 1 Asp 5	768
tti	t gc	a ga	c gt	t aa	a gc	a gto	c at	g aaa	_	t tc		a ac	t gc	a at	g ctt 	816

								(	)4sec	u.t	κt					
Phe	Ala	Asp	va1 260	Lys	Ala	۷al	Met	Lys 265	Asp	ser	Gly	Thr	Ala 270	Met	Leu	
ggt Gly	gtc Val	ggt Gly 275	gtt Val	tcc Ser	tca Ser	agt Ser	aaa Lys 280	aac Asn	cga Arg	gct Ala	gaa Glu	gaa Glu 285	gca Ala	gct Ala	gaa Glu	864
caa Gln	gca Ala 290	act Thr	ctt Leu	gct Ala	cct Pro	ttg Leu 295	att Ile	gga Gly	tca Ser	tca Ser	att Ile 300	caa Gln	tct Ser	gct Ala	aca Thr	912
ggt Gly 305	gtt val	gtt Val	tat Tyr	aat Asn	att Ile 310	acc Thr	gga Gly	ggg Gly	aag Lys	gac Asp 315	ata Ile	act Thr	cta Leu	caa Gln	gaa Glu 320	960
gtc Val	aac Asn	agg Arg	gtt Val	tct Ser 325	cag Gln	gtg Val	gta Val	aca Thr	agt Ser 330	ttg Leu	gca Ala	gat Asp	cca Pro	tca Ser 335	gca Ala	1008
aac Asn	att Ile	ata Ile	ttc Phe 340	ggg Gly	gca Ala	gtg Val	gta Val	gat Asp 345	gag Glu	aga Arg	tac Tyr	aac Asn	ggg Gly 350	Glu	att Ile	1056
)t	gtg Val	acc Thr 355	att Ile	gtt Val	gct Ala	act Thr	ggc Gly 360	ttt Phe	gcc Ala	cag Gln	tcg Ser	Phe 365	Gin	aaa Lys	tct Ser	1104
ctt Leu	ctt Leu 370	Ala	gac Asp	ccġ Pro	aaa Lys	gga Gly 375	gca Ala	aaa Lys	ctt Leu	gtt Val	gat Asp 380	Arg	aat Asn	caa Gln	gaa Glu	1152
cct Pro 385	Thr	caa Gln	cct Pro	ttg Leu	act Thr 390	ser	gcg	aga Arg	tct Ser	ttg Leu 395	i Thr	aca Thr	cct Pro	tct Ser	cct Pro 400	1200
gct Ala	ccg Pro	tct Ser	cgg Arg	tct Ser 405	' Arg	aaa Lys	cto Leu	tto Phe	ttt Phe 410	<b>?</b>	ι					1233

<210> 78

<211> 410

<212> PRT

213> Tagetes erecta

<400> 78

Met Ala Thr His Lys Leu Leu Gln Phe Thr Thr Asn Leu Pro Pro Ser 10 15

Ser Ser Ser Ile Ser Thr Gly Cys Ser Leu Ser Pro Phe Phe Leu Lys 20 25 30

Ser Ser Ser His Ser Pro Asn Pro Arg Arg His Arg Arg Ser Ala Val

Cys Cys Ser Phe Ala Ser Leu Asp Ser Ala Lys Ile Lys Val Val Gly 50 60

Val Gly Gly Gly Asn Asn Ala Val Asn Arg Met Ile Gly Ser Gly 65 70 75 Leu Gln Gly Val Asp Phe Tyr Ala Ile Asn Thr Asp Ser Gln Ala Leu 85 90 95 Leu Gln Ser Val Ala His Asn Pro Ile Gln Ile Gly Glu Leu Leu Thr 100 105 110 Arg Gly Leu Gly Thr Gly Gly Asn Pro Leu Leu Gly Glu Gln Ala Ala 115 120 125 Glu Glu Ser Lys Glu Ala Ile Gly Asn Ala Leu Lys Gly Ser Asp Leu 130 140 Val Phe Ile Thr Ala Gly Met Gly Gly Gly Thr Gly Ser Gly Ala Ala 145 150 150 b Val Val Ala Gln Ile Ala Lys Glu Ala Gly Tyr Leu Thr Val Gly 165 170 175 Val Val Thr Tyr Pro Phe Ser Phe Glu Gly Arg Lys Arg Ser Val Gln 180 185 Ala Leu Glu Ala Ile Glu Lys Leu Gln Lys Asn Val Asp Thr Leu Ile . 195 200 205 Val Ile Pro Asn Asp Arg Leu Leu Asp Ile Ala Asp Glu Asn Thr Pro 210 220 Leu Gln Asp Ala Phe Leu Leu Ala Asp Asp Val Leu Arg Gln Gly Val 225 230 235 240 Gln Gly Ile Ser Asp Ile Ile Thr Ile Pro Gly Leu Val Asn Val Asp 245 250 255 he Ala Asp Val Lys Ala Val Met Lys Asp Ser Gly Thr Ala Met Leu 260 265 270 Gly Val Gly Val Ser Ser Ser Lys Asn Arg Ala Glu Glu Ala Ala Glu 275 280 . 285 Gln Ala Thr Leu Ala Pro Leu Ile Gly Ser Ser Ile Gln Ser Ala Thr 290 295 300 Gly Val Val Tyr Asn Ile Thr Gly Gly Lys Asp Ile Thr Leu Gln Glu 305 310 315 Val Asn Arg Val Ser Gln Val Val Thr Ser Leu Ala Asp Pro Ser Ala 325 330

```
O4sequ.txt
Asn Ile Ile Phe Gly Ala Val Val Asp Glu Arg Tyr Asn Gly Glu Ile
340
345
His Val Thr Ile Val Ala Thr Gly Phe Ala Gln Ser Phe Gln Lys Ser
355 360 365
Leu Leu Ala Asp Pro Lys Gly Ala Lys Leu Val Asp Arg Asn Gln Glu
370 380
 Pro Thr Gln Pro Leu Thr Ser Ala Arg Ser Leu Thr Thr Pro Ser Pro 385 390 395
 Ala Pro Ser Arg Ser Arg Lys Leu Phe Phe 410
        79
 <210>
         33
  <211>
         DNA
  213> Künstliche Sequenz
  <220>
          primer_bind
  <221>
          (1)..(33)
  <222>
  <223>
                                                                                  33
   <400> 79
   gcatgctcta gaccttataa agatattttg tga
   <210>
           80
    211>
           33
     12>
           DNA
           Künstliche Sequenz
    <213>
    <220>
           primer_bind
    <221>
            (1)..(33)
    <222>
    <223>
```

<400> 80

gcatgcatct agaaatggtt cagtgtcaac cat

```
<210>
      81
<211>
       805
       DNA
<212>
       Nostoc sp. Strain PCC7120 .
<213>
<220>
<221> variation
<222> (1)..(805)
 <223>
 gcatgcatct agaaatggtt cagtgtcaac catcatctct gcattcagaa aaactggtgt
                                                                        60
   ttgtcatc gacaatcaga gatgataaaa atattaataa gggtatattt attgcctgct
                                                                       120
   atcttatt tttatgggca attagtttaa tcttattact ctcaatagat acatccataa
                                                                       180
 ttcataagag cttattaggt atagccatgc tttggcagac cttcttatat acaggtttat
                                                                       240
 ttattactgc tcatgatgcc atgcacggcg tagtttatcc caaaaatccc agaataaata
                                                                        300
  attttatagg taagctcact ctaatcttgt atggactact cccttataaa gatttattga
                                                                        360
  aaaaacattg gttacaccac ggacatcctg gtactgattt agaccctgat tattacaatg
                                                                        420
  gtcatcccca aaacttcttt ctttggtatc tacattttat gaagtcttat tggcgatgga
                                                                        480
  cgcaaatttt cggattagtg atgatttttc atggacttaa aaatctggtg catataccag
                                                                        540
  aaaataattt aattatattt tggatgatac cttctatttt aagttcagta caactatttt
                                                                        600
  attttggtac atttttgcct cataaaaagc tagaaggtgg ttatactaac ccccattgtg
                                                                        660
  cgcgcagtat cccattacct cttttttggt cttttgttac ttgttatcac ttcggctacc
                                                                        720
  acaaggaaca tcacgaatac cctcaacttc cttggtggaa attacctgaa gctcacaaaa
                                                                         780
                                                                         805
     tctttata aggtctagag catgc
          82
   <210>
   <211> 37
   <212> DNA
   <213> Künstliche Sequenz
   <220>
    <221> Primer
           (1)..(37)
    <222>
    <223>
```

<400> 82 gcgcatgcat ctagaaatga tccagttaga acaacca	37
<210> 83	
<211> 37	
<212> DNA	
<213> Künstliche Sequenz	
<220>	
<221> Primer	
<222> (1)(37)	
<223>	
<pre>&lt;400&gt; 83 gcgcatgctc tagactattt tgctttgtaa atttctg</pre>	37
gcgcatgete tagastas y	
<210> 84	
<211> 792 .	
<212> DNA	
<213> Nostoc punctiforme ATCC 29133	
<220>	
<221> misc_feature	
<222> (5)(775)	
223>	
<400> 84 gcgcatgcat ctagaaatga tccagttaga acaaccactc agtcatcaag caaaactgac	60
tccagtactg agaagtaaat ctcagtttaa ggggcttttc attgctattg tcattgttag	120
cgcatgggtc attagcctga gtttattact ttcccttgac atctcaaagc taaaattttg	180
gatgttattg cctgttatac tatggcaaac atttttatat acgggattat ttattacatc	240
tcatgatgcc atgcatggcg tagtatttcc ccaaaacacc aagattaatc atttgattgg	300
aacattgacc ctatcccttt atggtctttt accatatcaa aaactattga aaaaacattg	360
gttacaccac cacaatccag caagctcaat agacccggat tttcacaatg gtaaacacca	420
gttacaccac cacaatccag caagcicaat agacceggat sounds 5 o aagtttcttt gcttggtatt ttcattttat gaaaggttac tggagttggg ggcaaataat	480
aagtttettt gettggtatt tteattitat gaaaggttae tggggs-555 55 tgcgttgact attatttata actttgctaa atacatactc catateccaa gtgataatct	540
tgcgttgact attatttata actttgctaa atacatacte databasees 5 5	

aacttactt	t tgggtgctac	cctcgctttt	aagttcatta	caattattct	attttggtac	600
ttttttaco	cc catagtgaac	caataggggg	ttatgttcag	cctcattgtg	cccaaacaat	660
tagccgtc	ct atttggtggt	catttatcac	gtgctatcat	tttggctacc	acgaggaaca	720
tcacgaata	at cctcatattt	cttggtggca	gttaccagaa	atttacaaag	caaaatagtc	780
tagagcat	gc gc					792
<210> 8	5					
<211> 2						
<212> D						
<213> K	ünstliche Sequ	uenz				
<220>					·	
21> P	rimer					
<222> (	1)(26)					
<223>						
	5 ctg ctttaatgag	atatgc				26
-						
	36					
<211> 2						
	ONA					
<5T3> k	Künstliche Seq	luenz				
220>						
221> H	Orimer			•		
	(1)(27)					
<223>	(2) (2. )					
<400>						2"
ctcgagc <sup>-</sup>	ttg gacaatcag	t aaattga				27
<210>	87					
<211>	210					
<212>	DNA					
~21 <b>3</b> ~	Agrobacterium	tumefacien	<b>c</b>			

<220>	
<221> Terminator	
<222> (1)(210)	
<223>	
<400> 87 gtcgaccctg ctttaatgag atatgcgaga cgcctatgat cgcatgatat ttgctttcaa ttctgttgtg cacgttgtaa aaaacctgag catgtgtagc tcagatcctt accgccggtt tcggttcatt ctaatgaata tatcacccgt tactatcgta tttttatgaa taatattctc	60 120 180 210
cgttcaattt actgattgtc caagctcgag	210
<pre>&lt;210&gt; 88  11&gt; 1773  &lt;212&gt; DNA  &lt;213&gt; Petunia hybrida  &lt;220&gt; &lt;221&gt; Promotor &lt;222&gt; (1)(1773) &lt;223&gt;</pre>	
<400> 88 contracting togggaaaaa atgatgtggc	60
<pre>&lt;400&gt; 88 gagctctagc gcaatcttat gtggtacaaa tcttgattag tcgggaaaaa atgatgtggc gagctctagc gcaatcttat gtggtacaaa tcttgattag gcgttatgg gttgttgaag</pre>	120
cctacaaatg gttggaggat gggagatttg gctctatcta gagttatgtg gttgttgaag	180
tttggtta ctctctgctg tggtagttgg catatccaca ttgtctcctt ccacttttat	240
gacaattacg tgaaagttat gggttgtttt gtctattttt gtcgaggcct ttcttttcct tccaggttgt tgaagatggt ccaattcgat tagaataatg ttttgagctt tagcatattc	300
tccaggttgt tgaagatggt ccaattcgat tagaataatg cools s  tctctcgttt acacgattat agtaataatg atataggatg acagaagttg acacataaat	360
tetetegttt acaegattat aglaataatg acataggues news s t tttttattet etecatttae tttaateeaa ateteaceta eeetaaaett etttaatatg	420
tattcaatag tctatccgag taaattgtaa atttaacaac cattgataat attgacacct	480
actaacatat actagtaaag agaatattaa catggcacat ataatttgat gcaaaatgag	540
tatgatgaaa tttaaaccca aaatctcttg attttgacag tgtcaccttg acttgttaac	600
tatgatgaaa tttaaaccca daaccccetg debosyang 5 taataagtca tgttttagtg gcagaaagac aaactcatcc accaactgta tagcaataaa	660
aaatagaaga atcttcctga ggcaaagttt tggaaaaatt aagagtggct gagatttaat	720
aaatagaaga attiteetga ggeddag 100 05	780

ttcaacagga attagttcca cttaactttt aggttacgat acagtgctaa ttaaataact

Seite 129

780

		•		V 10 0 qui 1 u			
	taattgtatt	agatatttct	tgcacctaaa	aaatttaaaa	actgaaaaaa	ggtagcaatc	840
	aaaataaaca	aaaggacaaa	ataagtgaaa	ggtacagcca	ccaaccctgg	cggctcactg	900
	tttgttggtt	aaaacgtaga	cttacaccta	ccaaaatcta	caactaaaat	gaggcaataa	960
	tactttgccc	aaaattacca	agaaaagaaa	aagaaaggaa	tcccttaata	ttactctcct	1020
	ccatttcaca	ataaatatcc	tagtttgact	taaattagag	tttaaaaaat	gaaagacgac	1080
	ttttaaaact	tgtaatctaa	aataaatcat	agttaaatgt	gtggctataa	atcattgtat	1140
	taacggtaaa	gtggtaagtt	taaaagttaa	ttgttttcaa	atataaaatt	gtactatcat	1200
	tctttttgga	atggactaat	aagaaaacta	tgacatccat	tatggagcgg	agggagtatc	1260
	tccttttaac	aataaccttt	gtcccttcaa	ttcaattatc	agtatgcaaa	cattaaaaat	1320
	tattattgat	gttaagtacc	acatcatcct	taatgataga	atcatcgtag	aacgcttttc	1380
	caggcacaca	ttcaaactag	ttagaccagt	accacacatc	gaatattcca	gacttctttg	1440
۱	tgaatagt	: cgactacatt	ggataatgga	acttctcgaa	ttaacttcga	attagtcgag	1500
	caaaataa	tatatacgtc	gggtggaaaa	ctataaaatg	tttgacaaaa	atgtcaaatt	1560
	aatatatcaa	tctgcaacaa	ccttttcacc	ttgagaacac	agctgaaatt	ttttacaaag	1620
	gtagttggtg	g aagctagtca	gcgaatccca	ttaccttcca	ctctacctaa	cccccttcac	1680
	caacaacaaa	tttctgtaat	ttaaaaacta	gccaaaaaag	aactctcttt	tacaaagagc	1740
	caaagactca	a atctttactt	tcaagaaaag	ctt			1773
	<210> 89						
	<211> 29						
	<212> DN/	<b>1</b>					
		r nstliche Seq	แยกว				
	1225	iserrene seq	Ju Ci i L				
	220>						
		imer					
		)(29)					
	<223>	, (_0)					
	<400> 89 gagctctage	c gcaatcttat	gtggtacaa				_ 29
	<210> 90						

<211> 29 <212> DNA

<213> Künstliche Sequenz

```
<220>
<221> Primer
<222> (1)..(29)
<223>
<400> 90
                                                                      29
aagcttttct tgaaagtaaa gattgagtc
<210> 91
<211> 39
<212> DNA
<213> Künstliche Sequenz
<220>
<221> Primer
<222> (1)..(39)
<223>
<400> 91
                                                                      39
gcgcatgcat ctagaaatga atttttgtga taaaccagt
<210> 92
<211> 37
<212> DNA
  213> Künstliche Sequenz
 <220>
 <221> Primer
 <222> (1)..(37)
 <223>
 <400> 92
                                                                       37
 gcgcatgctc tagattacga attggttact gaattgt
 <210> 93
 <211> 819
```

<212> DNA <213> Nostoc punctiforme ATCC 29133 <220> misc\_feature <221> <222> (5)..(802)<223> <400> 93 gcgcatgcat ctagaaatga atttttgtga taaaccagtt agctattatg ttgcaataga 60 .120 qcaattaagt gctaaagaag atactgtttg ggggctggtg attgtcatag taattattag tctttgggta gctagtttgg cttttttact agctattaat tatgccaaag tcccaatttg 180 tgatacct attgcaatag titggcaaat gticctitat acagggctat tiattacigc 240 atgatgct atgcatgggt cagtttatcg taaaaatccc aaaattaata attttatcgg 300 360 ttcactagct gtagcgcttt acgctgtgtt tccatatcaa cagatgttaa agaatcattg cttacatcat cgtcatcctg ctagcgaagt tgacccagat tttcatgatg gtaagagaac 420 480 aaacgctatt ttctggtatc tccatttcat gatagaatac tccagttggc aacagttaat 540 agtactaact atcctattta atttagctaa atacgttttg cacatccatc aaataaatct catcttattt tggagtattc ctccaatttt aagttccatt caactgtttt atttcggaac 600 atttttgcct catcgagaac ccaagaaagg atatgtttat ccccattgca gccaaacaat 660 aaaattgcca acttttttgt catttatcgc ttgctaccac tttggttatc atgaagaaca 720 tcatgagtat ccccatgtac cttggtggca acttccatct gtatataagc agagagtatt 780 819 caacaattca gtaaccaatt cgtaatctag agcatgcgc 210> 94 11> 33 <212> DNA <213> Künstliche Sequenz <220> <221> Primer <222> (1)..(33) <223>

<400> 94 gcgcatgcat ctagaaatgg cgatcgccat tat

33

<210> 95	
<211> 32	
<212> DNA	
<213> Künstliche Sequenz	
<220>	
<221> primer	
<222> (1)(32)	
<223>	
<400> 95 gcatgctc tagatcacaa atttgattta ga	32
<210> 96	
<211> 720	
<212> DNA	
<213> Nodularia spumigena NSOR10	
<220>	
<221> CDS	
<222> (5)(703)	
<223> .	
100> 96 gc atg cat cta gaa atg gcg atc gcc att att agt ata tgg gct atc Met His Leu Glu Met Ala Ile Ala Ile Ile Ser Ile Trp Ala Ile 10 15	49
1 so the test case the age that tog	97
agc cta ggt ttg tta ctt tat att gat ata tcc caa ttc aag ttt tgg Ser Leu Gly Leu Leu Tyr Ile Asp Ile Ser Gln Phe Lys Phe Trp 25 30	
20	145
atg ttg tta ccg ctc ata ttt tgg caa aca ttt tta tat acg gga tta Met Leu Leu Pro Leu Ile Phe Trp Gln Thr Phe Leu Tyr Thr Gly Leu 40	•
33	193
ttt att aca gct cat gat gcc atg cat ggg gta gtt ttt ccc aaa aat Phe Ile Thr Ala His Asp Ala Met His Gly Val Val Phe Pro Lys Asn 60	
50	241
ccc aaa atc aac cat ttc att ggc tca ttg tgc ctg ttt ctt tat ggt Pro Lys Ile Asn His Phe Ile Gly Ser Leu Cys Leu Phe Leu Tyr Gly 70 75	-· <b>-</b>
65	289
ctt tta cct tat caa aaa ctt tta aaa aag cat tgg cta cat cac cat Seite 133	

04sequ.txt Leu Leu Pro Tyr Gln Lys Leu Leu Lys Lys His Trp Leu His His 90 95	
aat cca gcc agt gaa aca gat cca gat ttt cac aac ggg aag cag aaa  aat cca gcc agt gaa aca gat cca gat ttt cac aac ggg aag cag aaa  Asn Pro Ala Ser Glu Thr Asp Pro Asp Phe His Asn Gly Lys Gln Lys  105 110	337
100 and tag and tag	385
Asn Phe Phe Ata 11p Tyl Led 17 120 125	433
tta caa att atc aca tta atg att att tat aac tta cta aaa tat ata Leu Gln Ile Ile Thr Leu Met Ile Ile Tyr Asn Leu Leu Lys Tyr Ile 130 135	433
tgg cat ttt cca gag gat aat atg act tat ttt tgg gta gtt ccc tca Trp His Phe Pro Glu Asp Asn Met Thr Tyr Phe Trp Val Val Pro Ser 155	481
att tta agt tct tta caa tta ttt tat ttt gga act ttt cta ccc cac  Ile Leu Ser Ser Leu Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His  160  170  175	529
t gag cct gta gaa ggt tat aaa gag cct cat cgt tcc caa act att gag cct gta Glu Gly Tyr Lys Glu Pro His Arg Ser Gln Thr Ile 180 180	577
agc cgt ccc att tgg tgg tca ttt ata act tgt tac cat ttt ggt tat Ser Arg Pro Ile Trp Trp Ser Phe Ile Thr Cys Tyr His Phe Gly Tyr 195	625
cat tac gaa cat cat gaa tac ccc cat gtt cct tgg tgg caa tta cca His Tyr Glu His His Glu Tyr Pro His Val Pro Trp Trp Gln Leu Pro 210 215	673
gaa att tat aaa atg tct aaa tca aat ttg tgatctagag catgcgc Glu Ile Tyr Lys Met Ser Lys Ser Asn Leu 225 230	720
<210> 97	
<211> 233 .	
<212> PRT	
213> Nodularia spumigena NSOR10	

97 <400>

Met His Leu Glu Met Ala Ile Ala Ile Ile Ser Ile Trp Ala Ile Ser 10 15

Leu Gly Leu Leu Tyr Ile Asp Ile Ser Gln Phe Lys Phe Trp Met 25

Leu Leu Pro Leu Ile Phe Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe 45

Ile Thr Ala His Asp Ala Met His Gly Val Val Phe Pro Lys Asn Pro 50 60

Us Ile Asn His Phe Ile Gly Ser Leu Cys Leu Phe Leu Tyr Gly Leu 65 70 75 80 Leu Pro Tyr Gln Lys Leu Leu Lys Lys His Trp Leu His His Asn 85 90 95 Pro Ala Ser Glu Thr Asp Pro Asp Phe His Asn Gly Lys Gln Lys Asn 100 105 110 Phe Phe Ala Trp Tyr Leu Tyr Phe Met Lys Arg Tyr Trp Ser Trp Leu 115 120 125 Gln Ile Ile Thr Leu Met Ile Ile Tyr Asn Leu Leu Lys Tyr Ile Trp 130 135 140 His Phe Pro Glu Asp Asn Met Thr Tyr Phe Trp Val Val Pro Ser Ile 145 150 155 160 u Ser Ser Leu Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Ser 165 170 175 Glu Pro Val Glu Gly Tyr Lys Glu Pro His Arg Ser Gln Thr Ile Ser 180 185 190 Arg Pro Ile Trp Trp Ser Phe Ile Thr Cys Tyr His Phe Gly Tyr His 195 200 205 Tyr Glu His His Glu Tyr Pro His Val Pro Trp Trp Gln Leu Pro Glu 210 215 220 Ile Tyr Lys Met Ser Lys Ser Asn Leu 225 230

Promotoren zur Expression von Genen in Tagetes

Zusammenfassung

Die vorliegende Erfindung betrifft die Verwendung von Promotoren zur Expression, vorzugsweise zur blütenspezifischen Expression von Genen in Pflanzen der Gattung Tagetes, die genetisch veränderten Pflanzen der Gattung Tagetes sowie ein Verfahren zur Herstellung von biosynthetischen Produkten durch Kultivierung der genetisch veränderten Pflanzen.

# Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/EP04/008624

International filing date: 31 July 2004 (31.07.2004)

Document type: Certified copy of priority document

Document details: Country/Office: DE

Number: 10 2004 007 623.5

Filing date: 17 February 2004 (17.02.2004)

Date of receipt at the International Bureau: 24 January 2005 (24.01.2005)

Remark: Priority document submitted or transmitted to the International Bureau in

compliance with Rule 17.1(a) or (b)



# This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

# **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:			
BLACK BORDERS			
☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES			
☐ FADED TEXT OR DRAWING			
☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING			
☐ SKEWED/SLANTED IMAGES			
COLOR OR BLACK AND WHITE PHOTOGRAPHS			
GRAY SCALE DOCUMENTS			
☐ LINES OR MARKS ON ORIGINAL DOCUMENT			
☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY			
□ other:			

# IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.